



A11106 978138

U.S. Department
of Commerce

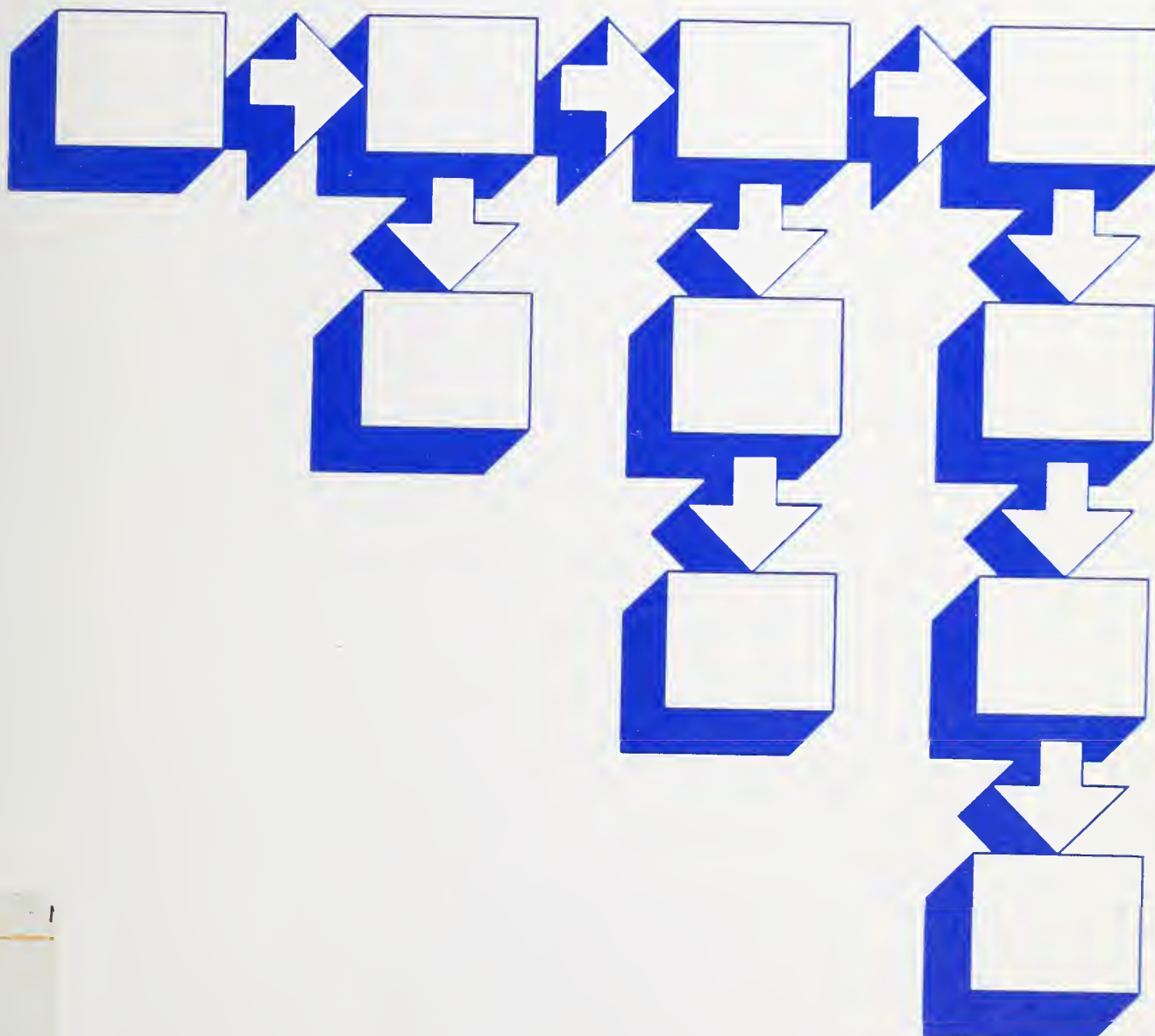
National Bureau
of Standards

Computer Science and Technology

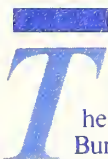
NBS Special Publication 500-123

Guide on Workload Forecasting

Helen Letmanyi



QC
100
U57
#500-123
1984



The National Bureau of Standards¹ was established by an act of Congress on March 3, 1901. The Bureau's overall goal is to strengthen and advance the nation's science and technology and facilitate their effective application for public benefit. To this end, the Bureau conducts research and provides: (1) a basis for the nation's physical measurement system, (2) scientific and technological services for industry and government, (3) a technical basis for equity in trade, and (4) technical services to promote public safety. The Bureau's technical work is performed by the National Measurement Laboratory, the National Engineering Laboratory, the Institute for Computer Sciences and Technology, and the Center for Materials Science.

The National Measurement Laboratory

Provides the national system of physical and chemical measurement; coordinates the system with measurement systems of other nations and furnishes essential services leading to accurate and uniform physical and chemical measurement throughout the Nation's scientific community, industry, and commerce; provides advisory and research services to other Government agencies; conducts physical and chemical research; develops, produces, and distributes Standard Reference Materials; and provides calibration services. The Laboratory consists of the following centers:

- Basic Standards²
- Radiation Research
- Chemical Physics
- Analytical Chemistry

The National Engineering Laboratory

Provides technology and technical services to the public and private sectors to address national needs and to solve national problems; conducts research in engineering and applied science in support of these efforts; builds and maintains competence in the necessary disciplines required to carry out this research and technical service; develops engineering data and measurement capabilities; provides engineering measurement traceability services; develops test methods and proposes engineering standards and code changes; develops and proposes new engineering practices; and develops and improves mechanisms to transfer results of its research to the ultimate user. The Laboratory consists of the following centers:

- Applied Mathematics
- Electronics and Electrical Engineering²
- Manufacturing Engineering
- Building Technology
- Fire Research
- Chemical Engineering²

The Institute for Computer Sciences and Technology

Conducts research and provides scientific and technical services to aid Federal agencies in the selection, acquisition, application, and use of computer technology to improve effectiveness and economy in Government operations in accordance with Public Law 89-306 (40 U.S.C. 759), relevant Executive Orders, and other directives; carries out this mission by managing the Federal Information Processing Standards Program, developing Federal ADP standards guidelines, and managing Federal participation in ADP voluntary standardization activities; provides scientific and technological advisory services and assistance to Federal agencies; and provides the technical foundation for computer-related policies of the Federal Government. The Institute consists of the following centers:

- Programming Science and Technology
- Computer Systems Engineering

The Center for Materials Science

Conducts research and provides measurements, data, standards, reference materials, quantitative understanding and other technical information fundamental to the processing, structure, properties and performance of materials; addresses the scientific basis for new advanced materials technologies; plans research around cross-country scientific themes such as nondestructive evaluation and phase diagram development; oversees Bureau-wide technical programs in nuclear reactor radiation research and nondestructive evaluation; and broadly disseminates generic technical information resulting from its programs. The Center consists of the following Divisions:

- Inorganic Materials
- Fracture and Deformation³
- Polymers
- Metallurgy
- Reactor Radiation

¹Headquarters and Laboratories at Gaithersburg, MD, unless otherwise noted; mailing address Gaithersburg, MD 20899.

²Some divisions within the center are located at Boulder, CO 80303.

³Located at Boulder, CO, with some elements at Gaithersburg, MD.

Computer Science and Technology

NBS
RESEARCH
INFORMATION
CENTER

QC100
.U57
#500-123
1984

NBS Special Publication 500-123

Guide on Workload Forecasting

Helen Letmanyi

Center for Programming Science and Technology
Institute for Computer Sciences and Technology
National Bureau of Standards
Gaithersburg, MD 20899



U.S. DEPARTMENT OF COMMERCE

Malcolm Baldrige, Secretary

National Bureau of Standards

Ernest Ambler, Director

Issued March 1985

Reports on Computer Science and Technology

The National Bureau of Standards has a special responsibility within the Federal Government for computer science and technology activities. The programs of the NBS Institute for Computer Sciences and Technology are designed to provide ADP standards, guidelines, and technical advisory services to improve the effectiveness of computer utilization in the Federal sector, and to perform appropriate research and development efforts as foundation for such activities and programs. This publication series will report these NBS efforts to the Federal computer community as well as to interested specialists in the academic and private sectors. Those wishing to receive notices of publications in this series should complete and return the form at the end of this publication.

Library of Congress Catalog Card Number: 85-600504

National Bureau of Standards Special Publication 500-123
Natl. Bur. Stand. (U.S.), Spec. Publ. 500-123, 68 pages (Mar. 1985)
CODEN: XNBSAV

U.S. GOVERNMENT PRINTING OFFICE
WASHINGTON: 1985

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402

CONTENTS

	Page
I. INTRODUCTION.....	2
A. Purpose.....	2
B. Background.....	3
C. Overview.....	3
II. OVERVIEW OF WORKLOAD FORECASTING TECHNIQUES.....	7
A. Forecasting in General.....	7
1. Forecast Spans.....	8
2. Data Patterns.....	8
2.1 Trend Data Pattern.....	8
2.2 Cyclical Data Pattern.....	10
2.3 Seasonal Data Pattern.....	10
2.4 Stationary Data Pattern.....	10
B. Criteria for Selecting Forecasting Techniques....	10
C. Forecasting Techniques.....	12
1. Simple Moving Averages.....	13
2. Exponential Smoothing.....	13
3. Classical Decomposition Method.....	14
4. Simple Linear Regression Analysis.....	15
5. Multiple Linear Regression Analysis.....	16
6. Box-Jenkins Method.....	16
III. WORKLOAD FORECASTING STEPS.....	18
STEP 1. Analyze Historical Data and Collect Future Requirements.....	18
1.1 Identify Forecasting Objectives.....	18
1.2 Organize Forecasting Team.....	18
1.3 Collect Data by Workload Levels.....	18
1.3.1 Level 1. Agency Mission.....	20
1.3.2 Level 2. Agency Functions.....	20
1.3.3 Level 3. Quantifiable Events....	21
1.3.4 Level 4. ADP Operations Performed.....	22
1.3.5 Level 5. Resources Consumed.....	22
1.4 Analyze Historical Data.....	22
1.4.1 Analyze Historical Data by Application Systems.....	23
1.4.2 Analyze Historical Data for the Total Workload.....	23
STEP 2. Select Forecasting Technique.....	24
STEP 3. Perform Workload Forecasting.....	25
3.1 Organizational Approach to Forecasting.	25
3.1.1 Forecast Level 3 Workload Requirements.....	29
3.1.2 Forecast Level 4 Workload Requirements.....	29
3.1.3 Forecast Level 5 Workload Requirements.....	29
3.2 Apply Forecasting Techniques.....	30
3.2.1 Forecasting with Simple Moving Averages.....	30

3.2.2 Forecasting with Exponential Smoothing.....	34
3.2.3 Forecasting with Classical Decomposition Method.....	35
3.2.4 Forecasting with Simple Linear Regression.....	43
3.2.5 Forecasting with Multiple Linear Regression.....	49
3.2.6 Forecasting with Box-Jenkins Method.....	53
STEP 4. Analyze Forecast Results.....	58
SUMMARY.....	60
APPENDIX A: SUGGESTED READINGS AND REFERENCES.....	61

GUIDE ON WORKLOAD FORECASTING

Helen Letmanyi

ABSTRACT

The purpose of this guide is to provide ADP managers and technical personnel with useful quantitative techniques for forecasting future workload requirements. It additionally provides a step-by-step approach to the forecasting process. Readers can then, in a timely manner, provide the computing resources needed to perform the user's workload at required service levels throughout the life-cycle of an ADP system. These techniques are described so that readers with little or no training in statistics should find them useful. However, this guide does not intend to give an exhaustive treatment of the techniques discussed. Readers requiring more information are referred to Appendix A ("Suggested Readings and References").

Key words: causal models; forecast spans; time-series models; workload forecasting techniques; workload levels; workload transition.

I. INTRODUCTION

A. Purpose

As a general planning tool, forecasting is a means by which the impact of today's decisions concerning future activities can be evaluated. As an ADP planning tool, the forecasting and evaluation of future workload requirements of an organization provides a basis for informed management decisions.

The purpose of this guide is to provide ADP managers and technical personnel with useful quantitative techniques for forecasting future workload requirements. It additionally provides a step-by-step approach to the forecasting process. Readers can then, in a timely manner, provide the computing resources needed to perform the user's workload at required service levels throughout the life-cycle of an ADP system. These techniques are described so that readers with little or no training in statistics should find them useful. However, this guide does not intend to give an exhaustive treatment of the techniques discussed. Readers requiring more information are referred to Appendix A ("Suggested Readings and References").

There are circumstances when the quantitative methods presented in this guide are either inappropriate or need to be applied with extreme caution. This is true in those situations where either historical data does not exist or where a trend is expected not to continue. In such an environment, qualitative, or non-quantitative, methods in lieu of or as a complement to those discussed in this guide may be appropriate.

No general guide of this kind can address every contingency; thus, specific decisions and actions in support of the workload forecasting process will vary from agency to agency. However, the workload forecast procedure described in this guide is applicable throughout the life-cycle of an ADP system, and should be an on-going process within an organization.

Portions of this guide are directed to technical staff, ADP management, and functional management. The Introduction, Overview, and the Overview of Workload Forecasting Techniques sections, as well as workload forecasting STEP 4 ("Analyze Forecast Results"), will be useful to functional management. Technical staff and ADP managers, who generally are responsible for forecasting the workload, should find the entire document useful.

B. Background

Forecasting the user's workload requirements plays an essential role in the life-cycle management of an ADP system. It significantly impacts the success that an organization has in providing the necessary computing resources in a timely manner.

A detailed description of the life-cycle of an ADP system is not within the scope of this guide. However, it is important to note that the forecasting process is an integral part of the life-cycle management of an ADP system. The life-cycle of an ADP system is generally defined as a collection of several phases, progressing from the requirements study to system operation. In general, an ADP life-cycle encompasses four main phases as shown in Figure 1.

Without an on-going workload forecast, the computing resources (system capacity) needed to perform the workload cannot effectively be estimated and planned. Figure 2 depicts the workload growth at a hypothetical organization. Assume, the organization of an ADP system took place in year 1 (point A). In year 4 (point B) the system - in its present configuration - reaches the saturation point. At this point, if possible, the system needs to be tuned/upgraded in order to handle the workload growth. Let's assume, that the system reaches its maximum practical capacity in year 8 (point D), that is, no additional tuning will provide better service. From year 6 (point C), in the absence of a capacity planning (including forecasting) activity, this organization would not be able to provide the needed resources to fulfill the organization's workload requirements. This could result in it not being able to support the organization's mission.

C. Overview

This section provides a brief outline of this guide. SECTION II. Description of Workload Forecasting Techniques gives a brief description of those forecasting techniques which are most widely used. Also, several criteria for selecting a given technique are discussed.

SECTION III. Workload Forecasting Steps describes a step-by-step approach to the workload forecasting process. The following four procedural steps are identified:

STEP 1. Analyze Historical Data and Collect Future Requirements. This step discusses the importance of the participation of different levels of users in the historical data collection process, and in the future

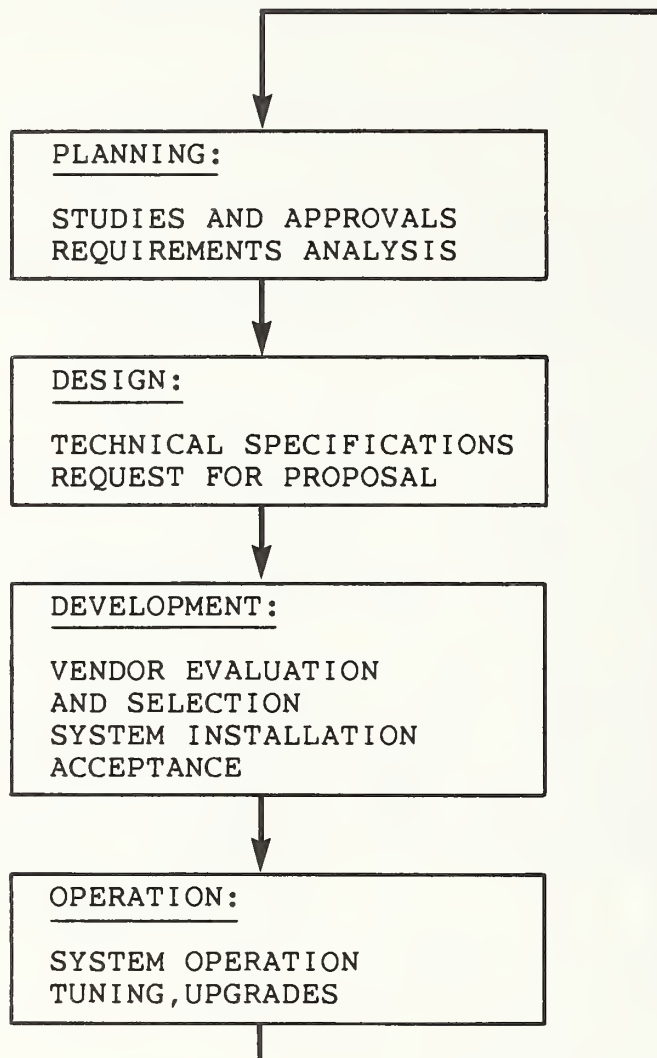


FIGURE 1. ADP SYSTEM LIFE CYCLE

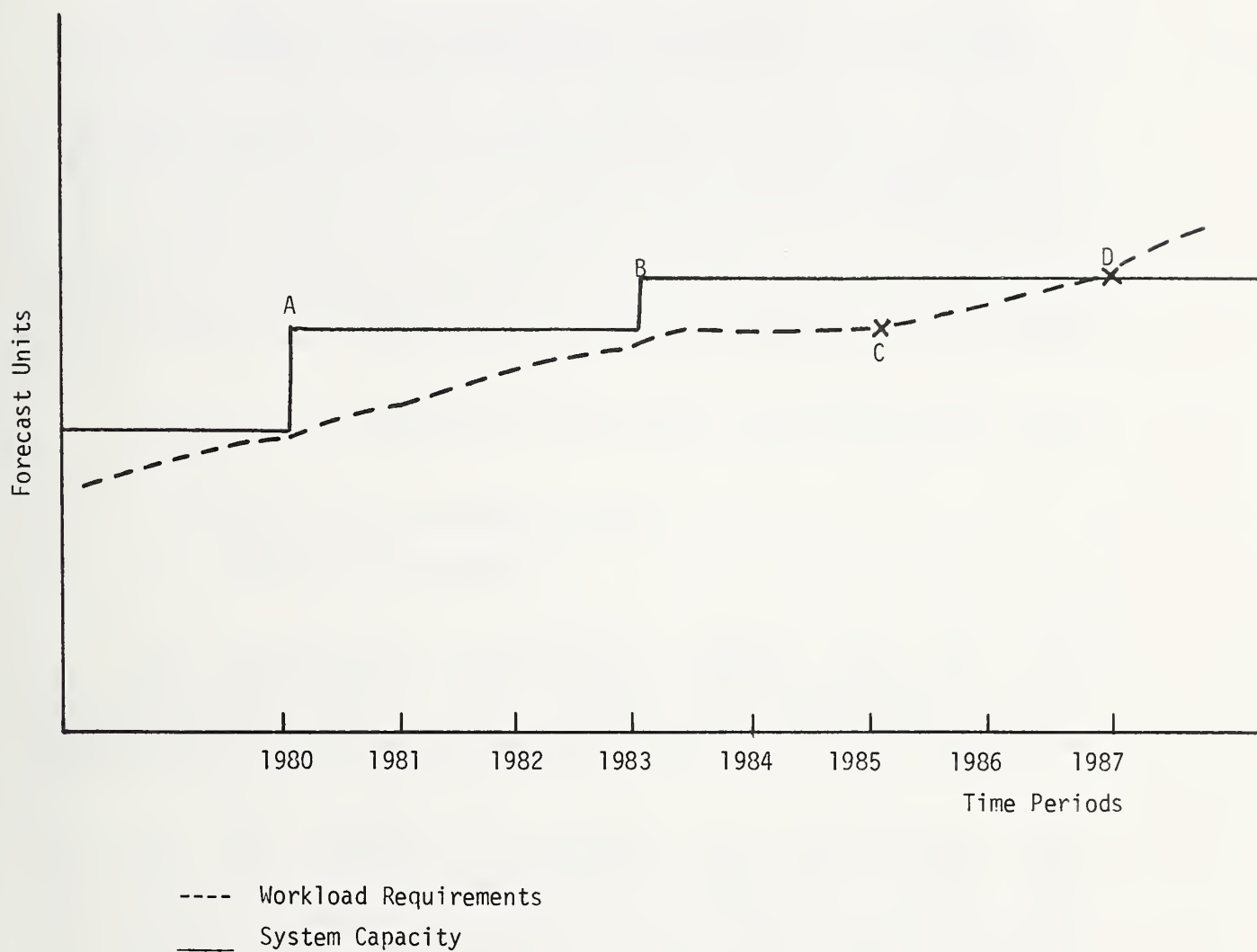


FIGURE 2. WORKLOAD REQUIREMENTS VERSUS SYSTEM CAPACITY

requirements identification procedure.

STEP 2. Select Forecasting Technique. This step identifies the criteria to be met in order to select a given technique.

STEP 3. Perform Workload Forecast. This step gives a practical example of how to apply each of the forecasting techniques discussed in the Overview of Workload Forecasting Techniques section of this guide.

STEP 4. Analyze Forecast Results. This step discusses the importance of analyzing the forecast results before any actions are taken.

II. OVERVIEW OF WORKLOAD FORECASTING TECHNIQUES

A. Forecasting in General

The distinction between planning and forecasting is often confused. For the purpose of this guide, forecasting is defined as a planning tool to predict what will "actually" happen or to identify possible outcomes. Planning is defined as a process that formulates objectives, and identifies the actions to be taken to achieve these objectives in light of the agency's mission. With every forecasting situation, in varying degrees, uncertainties are associated. Forecasting results should not be treated as absolutes. Forecasts need to be carefully evaluated, before decisions can be made based on their results. The terms forecasting and predicting are used interchangeably in this guide.

In general, forecasting techniques can be divided into two groups:

1. quantitative methods, and
2. qualitative methods.

The quantitative forecasting methods rely on the existence of historical data and on those variables that can be quantified. With this method of forecasting, those variables that are non-quantifiable are often ignored or given very little consideration.

For the qualitative methods of forecasting, the forecasts are based on judgements and expert opinions and usually with no direct reliance on historical data. The qualitative methods are especially useful for long-range forecasting, when no historical data is available or a trend is expected to change significantly. With these methods, it is usually assumed that forecasts can be made by analyzing the probable outcomes of cause-effect relations in a given situation under certain conditions. The Delphi Method [LI75] is one of the most widely used qualitative forecasting techniques. This method of forecasting takes the best judgement estimates from individuals determined to be expert in a given area. The refined consensus of these estimates becomes the forecast.

In this guide, only those quantitative techniques will be discussed which are most widely used in the workload forecasting process.

1. Forecast Spans

The forecast span is defined in this guide as the time horizon for which forecasting is to be performed. In this guide three forecast spans are identified:

1. Short-range forecast span. The short-range forecast span in the workload forecasting process is defined to be in the range of from one week to two years.

2. Medium-range forecast span. The medium-range forecast span in this guide is defined as a time horizon between two to five years.

3. Long-range forecast span. The long-range forecast span in the workload forecasting process is defined as a time horizon of five or more years.

2. Data Patterns

With the exception of the Box-Jenkins method described below, the workload forecasting techniques discussed in this guide assume that a certain underlying pattern can be identified in the historical data. In general, four basic patterns can be observed in historical data:

1. Trend Data Pattern;
2. Cyclical Data Pattern;
3. Seasonal Data Pattern;
4. Stationary Data Pattern.

It should be noted, that these patterns are not mutually exclusive; often they can be observed together. For predicting the workload it is sometimes desirable to separate the effects of each of the pattern components on the forecast value. This can be done by decomposition techniques, such as the Classical Decomposition Method (see SECTION III.).

2.1. Trend Data Pattern

A trend data pattern refers to a change where the historical workload increases or decreases over a longer period of time. In practice, the workload has the tendency to increase due to such factors as increase in the number of transactions and increase in program development. Figure 3

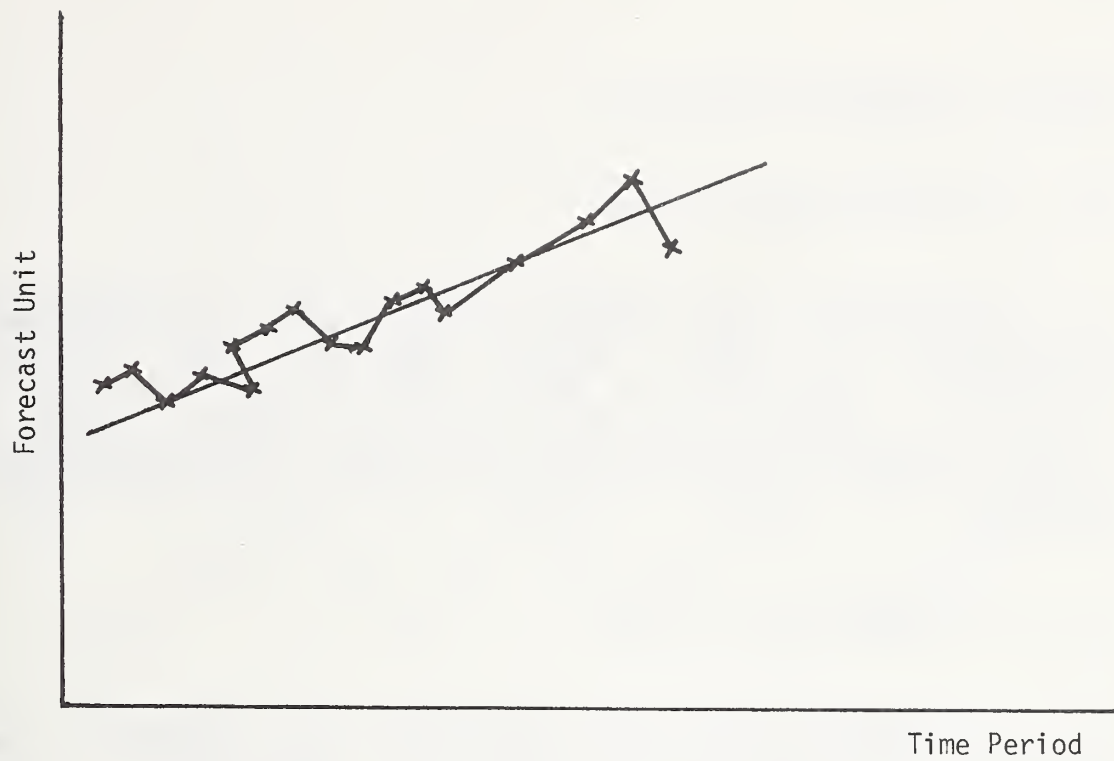


FIGURE 3. TREND DATA PATTERN



FIGURE 4. CYCLICAL DATA PATTERN

shows a trend data pattern.

2.2. Cyclical Data Pattern

A cyclical data pattern refers to the long range fluctuations in the workload requirements. The length of the pattern is usually longer than one year, and the fluctuation (pattern) is not repeated at identical time intervals. Cyclical data patterns are usually caused by factors, such as introduction of new application systems and elimination of others. Figure 4 depicts a cyclical data pattern.

2.3. Seasonal Data Pattern

A seasonal data pattern is similar to a cyclical data pattern except that the fluctuation is caused by some seasonal factor and the fluctuation is repeated at almost identical time intervals. That is, the seasonal pattern repeats itself usually within a year's interval. Seasonal data patterns can be found at many organizations where certain tasks need to be performed daily, weekly, monthly, etc., and can cause seasonal patterns in the workload data. Figure 5 shows a seasonal data pattern.

2.4. Stationary Data Pattern

Stationary data pattern refers to a series of observations that do not exhibit any systematic increase or decrease. The time series behaves as random fluctuations about a mean value and no trend can be identified. Figure 6 is an example of a stationary data pattern.

B. Criteria for Selecting Forecasting Techniques

In general, the amount of effort associated with using a forecasting technique should be approximately proportional to the benefit/criticality of the forecast to the organization as a whole. There are four main criteria in choosing forecasting techniques:

1. Data requirements and pattern of the data. Most of the quantitative forecasting techniques require the availability of reliable historical data. The success of a forecast process is significantly affected by the selection of the technique that can most

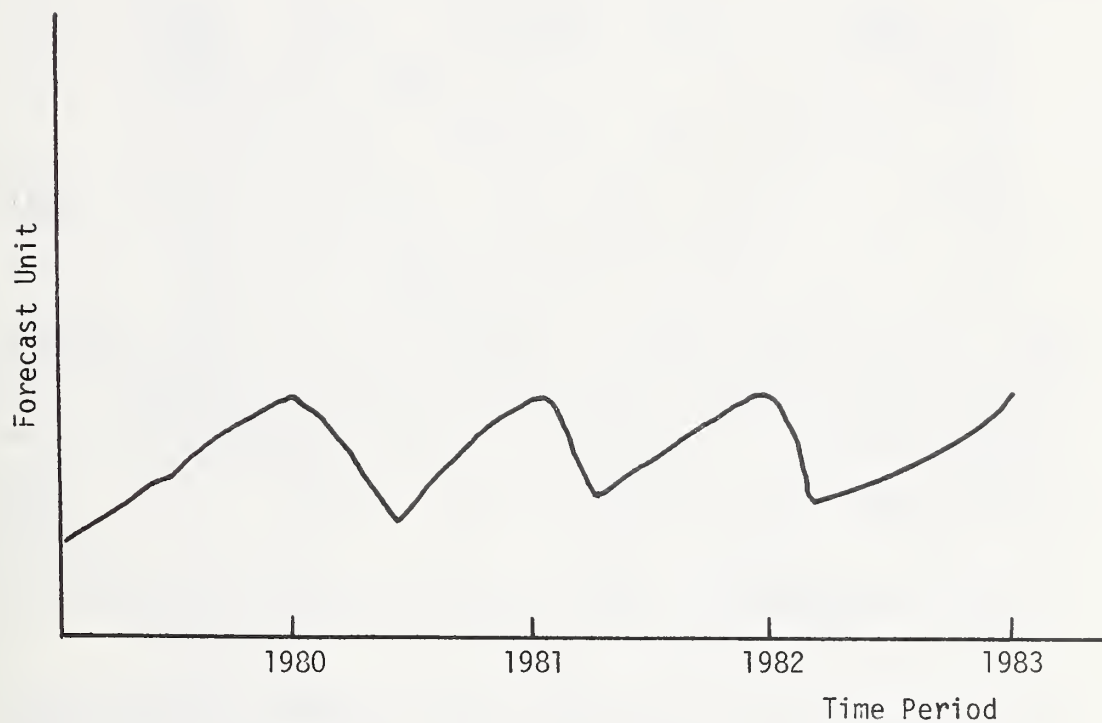


FIGURE 5. SEASONAL DATA PATTERN

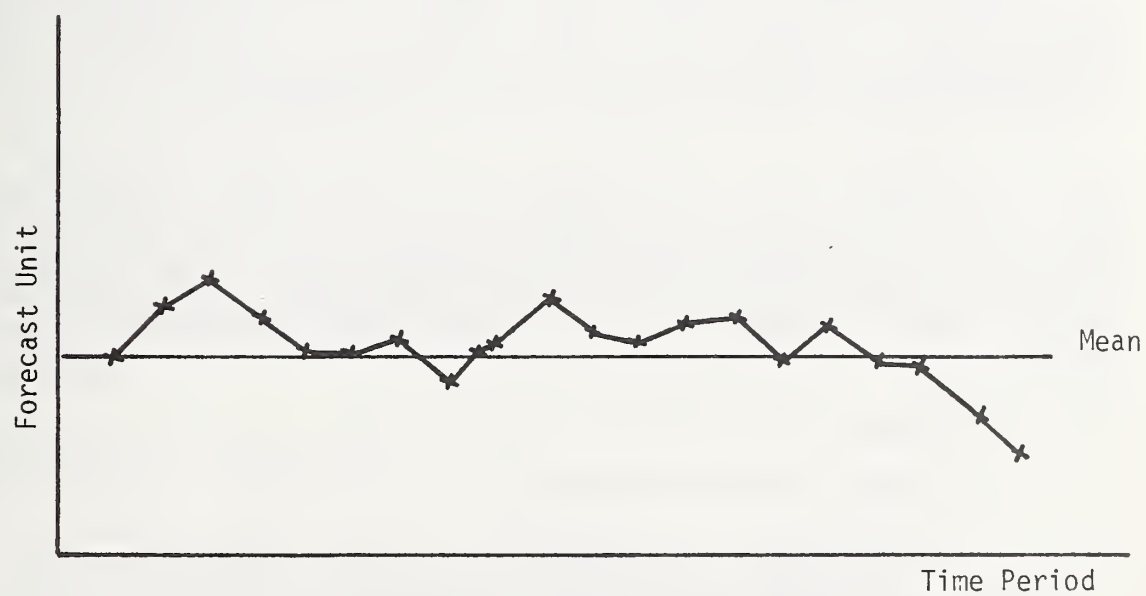


FIGURE 6. HORIZONTAL (STATIONARY) DATA PATTERN

effectively handle a given pattern type.

2. Accuracy versus cost of using the technique. In general, the more sophisticated the technique and the more factors considered, the more accurate and the more costly are the forecasts compared to those derived from simpler, less costly methods. However, when the reliability of the data is questionable or subjective, the usage of sophisticated techniques will not add to the accuracy of the forecast results.

3. Forecast span. Some forecasting techniques, such as simple moving averages are appropriate only for short-range forecasting. Others, like regression analysis are suitable for short- to long-range forecasting.

4. Management's confidence in the technique. Since management has the final authority on the course of actions to be taken, it cannot be emphasized enough that the communicability of the forecast results and management's confidence in the technique have a major impact on whether the results will be used.

C. Forecasting Techniques

This section discusses six quantitative techniques that are most commonly used for forecasting future workload requirements. Following a brief general description, the techniques are viewed from the perspective of the first three selection criteria presented in Section B.

"Management's confidence in the technique" as one of the criteria for the selection of a forecasting technique will not be further discussed, since this criterion is dependent on the background and preferences of the manager involved in the decision making process. Some managers might feel confident with a given technique, while others might not.

The techniques to be examined are:

1. Simple moving averages;
2. Exponential smoothing;
3. Classical decomposition method;

4. Simple linear regression analysis;
5. Multiple linear regression analysis; and
6. Box-Jenkins method.

1. Simple Moving Averages

In this method of forecasting, historical data is used to compute the forecast value. The average of the known observations (values) becomes the forecast for the next time period. When this forecast value then becomes a known value (an actual observation) the average is then recalculated for the next period. Hence, this method is called moving averages. Moving average methods are also referred to as "smoothing techniques," where extreme values of historical data are smoothed by averaging. The smoothing effect of this technique becomes greater with the increase in the number of observations included in the averaging.

Data requirements and pattern of the data. Historical data is required over several time periods (20-30), depending on the smoothing effect desired. The number of time periods selected should cover any periods with random fluctuations, if the effect of the extreme values is to be evaluated. The basic pattern of the data should be nearly stationary and this same pattern is assumed for the future.

Accuracy versus cost of using the technique. The advantages of using this method are that it is simple and inexpensive. However, simple moving averages are limited to one time period at a time into the future. If the pattern of the historical data shows only little variation and the change is nearly stationary, the confidence level using this technique is usually high.

Forecast span. The forecast span is determined by the interval of the time periods of the historical data. For example, if the historical data is available for every three months, then the forecast can be made for three months into the future. This technique is mostly used for short-range forecasting.

2. Exponential Smoothing

The underlying theory of exponential smoothing is the same as for moving averages: historical data provides the average which becomes the forecast value and which in turn becomes a new known value. The difference is that more recent observations are given exponentially increasing

smoothing weights. This is based on the assumption that more recent observations provide a better indication of what will happen in the future. The smoothing weight is determined to be between zero and one.

Data requirements and pattern of the data. In order to calculate the initial forecast value, historical data is needed over several time periods (20-30) and the smoothing weight must be determined. For subsequent forecasts, no historical data are needed. The only data required are the most recent observed value, the most recent forecast and the value for the smoothing weight. The pattern of historical data should be stationary and with no major changes expected in the future.

Accuracy versus cost of using the technique. The major advantages of using this technique are the low costs associated with the development of the forecast, and its ease of use. It is more accurate than the simple moving averages. However, it also requires a stationary pattern in the data.

Forecast span. As with simple moving averages, the forecast span is determined by the intervals (e.g., weekly, monthly) of the historical data.

3. Classical Decomposition Method

In the workload forecasting process, it is sometimes desirable to distinguish the individual pattern components from the underlying pattern in the historical data, i.e., to identify what part of the increase in the requirements is caused by seasonal, cyclical, or random fluctuations (about the trend), and what portion represents an overall increase (trend). The identification of the effect of the pattern components can provide valuable information to the operation personnel in the current scheduling process, in addition to forecasting. The major advantages of this method are that it is easy to use and interpret.

Data requirements and pattern of the data. For determining the different pattern components for a given set of time series data, historical data is needed over several seasonal/cyclical periods (6-7). Time series data is defined as historical data presented in uniform terms and covering a specified period of time. As noted above, this method can handle time series data with trend, seasonal, and cyclical components.

Accuracy versus cost of using this technique. A relatively accurate forecast can be achieved if sufficient historical data is available (e.g., covering six seasons/cycles). The development of this technique is expensive, but once the

model is developed, the application of the method is inexpensive.

Forecast span. This method of forecasting can be used from short- to long-range forecasting, although, it is most applicable for short-range.

4. Simple Linear Regression Analysis

Simple regression analysis is used to determine the linear relationship between two variables. That is, a mathematical relationship is determined between the dependent variable and the independent variable. This relationship is expressed in the regression coefficient of the regression equation. The statistical method of least squares is used to determine the equation of the relationship using historical data. The future values of the dependent variable can then be extrapolated or calculated using the equation determined by the least squares method.

Data requirements and pattern of the data. A large number of accurate historical data points (30-40) for both dependent and independent variables is required to determine the linearity of the relationship. In contrast with forecasting techniques discussed in the previous sections, this technique not only assumes that an underlying pattern exists in the historical data, but also that the basic pattern is linear. That is, if the historical data is plotted, the data points would fall along a straight line.

Accuracy versus cost of using the technique. This technique allows for testing of the statistical significance of the regression equation and of the individual coefficients, and obtaining bounds for the predicted values. High accuracy can be expected when no changes are anticipated to occur which may modify the linear relationship of the variables. The cost of developing the model for this technique can be expensive if the data is not readily available.

Forecast span. Simple regression methods can be used for any forecast span as long as a linear relationship exists between the variables and the same relationship can be expected for the future.

5. Multiple Linear Regression Analysis

This forecast method is similar to simple regression analysis. However, with this method, the linear relationship is defined among several independent variables to the dependent variable. This relationship is expressed in the regression coefficients of the regression equation. Using observations of historical data for the variables, correlation coefficients can be obtained by least squares method for the independent variables. The future values of the dependent variable can be computed (if future estimates for independent variables are available) or extrapolated using the equation determined by the least squares method.

Data requirements and pattern of the data. To successfully apply this technique, a large number of historical data points (30-40) is required to assure the linear relationship of the independent variables to the dependent variable. This technique can handle any kind of pattern as long as a linear relationship exists between the dependent and independent variables.

Accuracy versus cost of using the technique. This technique allows for testing of the statistical significance of the regression equation and of the individual coefficients, and obtaining bounds for the predicted values. These measures enable the forecaster to achieve a relatively high degree of accuracy. Collecting the required large amount of historical data to develop the initial regression equation can be an expensive undertaking, but forecasting the future values of the dependent variable using the regression equation is inexpensive.

Forecast span. There is no limit on the forecast span.

6. Box-Jenkins Method

In cases where the pattern of the historical data is readily apparent, methods discussed in the previous sections are generally less expensive to use and more easily understood than the more sophisticated Box-Jenkins method.

This method is a relatively recent quantitative forecasting technique to handle those complex time series data where the pattern is not easily identifiable. This forecasting method is based on pattern identification. It identifies the pattern in the historical data and uses this pattern as the basis for forecasting future requirements. Pattern recognition is accomplished by computing the autocorrelations among values of the same variable. The autocorrelation is a measure of the dependence among values of the same variable at different time periods. High

autocorrelation indicates seasonal and/or cyclical patterns. Like forecasting techniques discussed in the previous sections, the Box-Jenkins method forecasts a single variable. This method is a powerful forecasting tool where the variable to be forecasted is a complex function of a variety of factors.

The recognition of the pattern in the data is accomplished by multiple iterations of the data, with each iteration providing information that allows the next iteration to more closely approximate the actual pattern of the data.

Data requirements and pattern of the data. Since this technique is developed to handle complex time series data and/or where the basic pattern is not apparent, the only requirement is historical data over several time periods (covering 6-7 seasonal/cyclical periods).

Accuracy versus cost of using the technique. This technique is considered by many experts in the field as one of the most powerful and accurate forecasting techniques available today. However, it is complex and expensive to use in comparison to other techniques discussed.

Forecast span. This technique is mostly used for short- to medium-range forecasting.

III. WORKLOAD FORECASTING STEPS

STEP 1. Analyze Historical Data and Collect Future Requirements

As indicated in the previous sections, workload forecasting plays a central role in the life-cycle management of an ADP system and should be an on-going process at an organization. One of the basic criteria for successful forecasting is the availability of reliable historical data. The historical data requirements are mostly determined by the degree of criticality associated with the effect of the decisions being made, based on the forecast results.

1.1 Identify Forecasting Objectives

The forecasting objectives must be identified by the decision makers, if the forecasting process is to succeed. The forecasting objectives might include the determination of items such as: the criticality of the decisions to be made based on the forecast results, level of effort to be expended, desired forecast span, etc.

1.2 Organize Forecasting Team

A forecasting team will vary in size from agency to agency and from one forecasting activity to another. There are several factors that determine the size and composition of the forecasting team. These factors include the variety of existing and new functions/application systems, and whether historical data is available or needs to be collected. The forecasting team should include those personnel who are knowledgeable about every aspect of the workload description at a given workload level. The decision makers and personnel responsible for the planning at the organization should also be included in the forecasting team.

1.3 Collect Data by Workload Levels

The workload forecasting process requires input from several sources. At this point it seems appropriate to describe an agency's workload on different levels. Figure 7 depicts the different workload levels at a given organization.

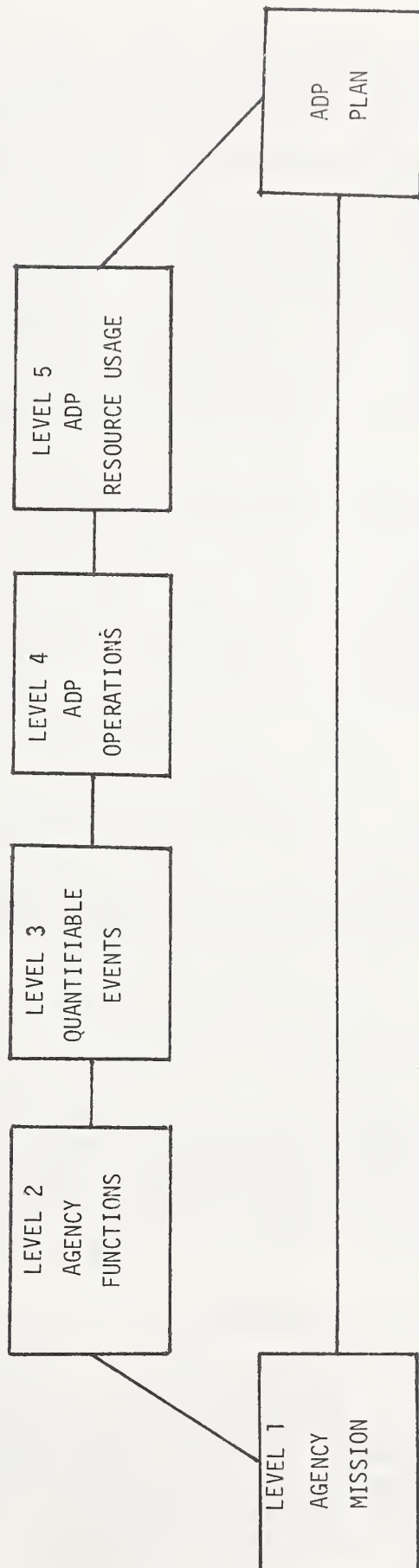


FIGURE 7. WORKLOAD TRANSLATION PROCESS

Each of these workload levels can be associated with different levels of management and/or users. As described in STEP 3 ("Perform Workload Forecast"), this description of the workload by levels provides the necessary input for translating mission requirements into ADP resource needs.

On the first four levels the workload is described in user oriented terms. In fact, it is almost imperative that these measures are used when forecasting workload requirements. These user oriented or functional measures can then be translated into hardware/software requirements whether the workload be performed on micros, mainframes, etc. To collect data and to perform forecasts for new functions/applications with no history, initial estimates can be based on similar operational applications.

The functional users (Level 3) usually can identify to the forecaster performing the survey, those personnel to be contacted for obtaining workload data on different workload levels. After identifying the personnel responsible for providing the workload data for the different workload levels, the forecaster should organize teams by functional areas. Also, the personnel performing the planning function should be included in the team. It is important to note, that before the actual data collection takes place on the different workload levels the variables/forecast units must be identified, which is not always trivial. This is the main reason why the participation of the personnel knowledgeable in every aspect of the workload description on the given workload level is essential to the success of the workload forecasting process. Several criteria need to be considered for selecting the variables/forecast units, for example, objective of the forecast, data availability and accuracy, etc.

1.3.1 Level 1. Agency Mission

The agency's mission is dictated by national interests and needs. The identification of the agency's mission assumes that an agency has a definite structure based on goals and objectives. The agency's mission takes form in the agency functions.

1.3.2 Level 2. Agency Functions

On this level, the workload is described by the functions necessary to support the agency's mission. These might include, for example, payroll, engineering, inventory, budget, transportation, etc. On this level, the workload can be associated with the top management. They are usually the decision makers, and the sources of information which

enable the users to describe (and forecast) their workload requirements on the next level.

The data that can be obtained on this level includes such important information as:

- type of functions to be performed,
- budget ceilings,
- frequency of reports generated,
- changes in number of employees.

This data and other agency-dependent information and decisions should be obtained to the extent that they impact future computing resource needs.

While collecting information on past, present, and future workloads, the agency functions should be evaluated in relation to the agency mission. This is done to see if any new functions are needed to support the agency's mission, and whether the present functions need to be continued into the future.

1.3.3 Level 3. Quantifiable Events

On this level, the workload can be described by those quantifiable events that can be associated with a given agency function. The quantifiable events are also referred to as "Natural Forecast Units" (NFU's). For example, the number of paychecks are the quantifiable events associated with the payroll function; the number of student loans are associated with the loan function. On this level, the workload description can be associated with functional users, such as personnel preparing the payroll. Information from these functional users is then provided to the direct users for describing their workload on the next level (Level 4). The functional users can provide the relationship between quantifiable events and agency functions. These are the users that can usually provide information such as:

- number of paychecks produced,
- number and type of transactions processed,
- frequency and schedule of production runs,
- number of loans processed.

1.3.4 Level 4. ADP Operations Performed

The workload on this level can be described in terms of ADP operations which must be performed in order to produce the given quantifiable events which support an agency function and ultimately a given agency mission. These users usually are those personnel who provide the programming and other support to the functional user. These users can provide the relationship between ADP operations and quantifiable events. The data that can be obtained on this level includes information such as:

number of records sorted,
number of records updated,
number of database queries,

for each quantifiable event.

1.3.5 Level 5. Resources Consumed

On this level, the workload can be described by the computing resources it consumes. For example, a job/jobstep, or a transaction processed, can be described by a set of resource demands. This information is usually provided by the operations personnel through the use of accounting logs. The resource usage data can usually be obtained from accounting logs. If the data is not readily available, the collection of the historical resource usage data by applications/application systems can be a lengthy procedure. Those organizations where the workload forecasting is an on-going process only need to update the resource usage information. The resource usage data is usually collected with the help of operations personnel.

1.4 Analyze Historical Data

Having collected the workload from the different levels, the forecaster needs to analyze the data. The historical data should be analyzed in two stages:

1. by application systems; and
2. by the total workload.

1.4.1 Analyze Historical Data by Application Systems

The analysis of the historical data by application systems can reveal important information necessary to build a model which will be used to forecast a given application system. The historical data collected from the different levels of users can be plotted using simple histograms and scatter diagrams. These simple techniques can provide the forecaster valuable insight into relationships among different levels of workload, such as:

- the effect of budget change on the number of quantifiable events,

- the effect of the increase in the number of quantifiable events on the volume of an associated ADP operation performed,

- change in resource requirements due to variations in the number of quantifiable events,

- how the rates of growth compare on the different workload levels,

- the pattern of the historical data for the application systems by workload levels in terms of some identifiable unit of measure.

1.4.2 Analyze Historical Data for the Total Workload

In order to determine the total workload requirements at a given organization, the aggregate workload requirements for all the application systems need to be obtained. The most frequently used unit for determining the total workload requirements is in terms of system accounting units (SAU's). The SAU can be found on most third-generation computers; examples of such units are CRU, SUP, SRU. In some cases, where the workload is homogeneous, the total workload requirements might be expressed in terms of functional measures such as the number of transactions processed.

Simple analysis tools, such as graphs, can be used to plot the total historical workload requirements against time periods (weekly, monthly, etc.) to observe the behavior of the total workload requirements at an organization.

The analysis of the graphs provides valuable information to the forecaster for identifying:

- the pattern of workload requirements in terms of resource usage, and

the underlying trend of the historical data.

Also, the graph for the total workload requirements can be used in conjunction with the graphs obtained for the application systems by workload levels to obtain information such as:

which of the applications/application systems are causing a certain pattern in the resource consumption,

what are the effects of the workload composition changes in the resource consumption.

This information can be useful for technical personnel involved in benchmark construction and capacity planning on the present system.

STEP 2. Select Forecasting Technique

As described in the Overview of Workload Forecasting Techniques section, there are several criteria to be considered in selecting a forecasting technique for a given situation including:

1. data availability and the pattern of the historical data;
2. accuracy versus cost of using the technique;
3. desired forecast span.

The analysis of the historical data and the future requirements (performed in STEP 1) provides a critical input for selecting a given forecasting technique. This analysis reveals the identifiable pattern of the historical data. Also, the availability of the future requirements gives a good indication whether causal models are appropriate or time-series analysis methods should be used.

As mentioned earlier, management's confidence in the forecasting technique to be selected plays an essential role in the success of the forecasting process. Management should be aware of the accuracy that can possibly be achieved, and also, the limitations of a technique intended to be employed.

It is important to note that the more sophisticated techniques do not necessarily lead to more accurate results. In fact, if the validity of the historical data is

questionable or the uncertainties associated with future requirements are too great, the use of complex techniques would not benefit the organization as a whole.

In addition to the criteria identified in Table 1, there are several other factors to be considered when selecting a forecasting technique for a given situation.

Organizations where no formal forecasting is in existence should always start with the simplest technique appropriate and progress to the more complex as necessary.

The criticality of the forecast results should also be taken into consideration when selecting forecasting techniques, that is, what the effect of the forecast results will be on the decision making.

STEP 3. Perform Workload Forecasting

At this point, it is assumed, the data are collected, organized, and analyzed. Also, the decisions have been made concerning the criticality of the forecast results, which determines the required accuracy and the amount of resources to be expended on the forecasting effort. The forecast unit(s) and the required forecast span should also have been determined.

Before applying any given model for forecasting workload requirements, the model derived through a forecasting technique needs to be validated. A way to perform this validation is by using only part of the historical data (if sufficient) to build the model, and exercise the model on the remaining data to see how the forecast values compare to the actual values.

In most cases, short- and medium-range forecasting quantitative techniques can be used. However, for long-range forecasting, in addition to quantitative techniques, qualitative evaluation of the forecast results should be performed through expert opinions and judgements.

3.1 Organizational Approach to Forecasting

The application of the forecasting techniques does not depend on the kind of data (e.g.; resource usage or ADP operations) used. In some cases the workload forecast is performed by using data from only one workload level, in others by using data from two different levels. The usage of the information from more than one workload level provides a tool by which an organization's missions/plans can be translated into processing demands expressed in terms

CRITERIA TECHNIQUE	DATA PATTERN	COMPLEXITY	FORECAST SPAN	RELATIVE COST
SIMPLE MOVING AVERAGES	STATIONARY	SIMPLE	SHORT-RANGE	LOW
EXPONENTIAL SMOOTHING	STATIONARY	SIMPLE	SHORT-RANGE	LOW
CLASSICAL DECOMPOSITION	SEASONAL/TREND/ CYCLICAL	SIMPLE	SHORT-RANGE	MEDIUM
SIMPLE LINEAR REGRESSION	LINEAR/TREND	SIMPLE	SHORT-TO-LONG RANGE	MEDIUM
MULTIPLE LINEAR REGRESSION	LINEAR/TREND	MODERATE	SHORT-TO-LONG RANGE	HIGH
BOX-JENKINS METHOD	ANY PATTERN	HIGH	SHORT-TO-MEDIUM RANGE	HIGH

TABLE 1. CRITERIA FOR SELECTING FORECASTING TECHNIQUES

of ADP operations/resource demands. These resource demands can then be used as input to the agency's ADP planning process to determine the actions (e.g., tuning, upgrade, replacement, etc.) to be taken to fulfill the agency's mission requirements. Figure 7 depicts this translation procedure (organizational approach) in the workload forecasting process.

The ultimate goal of an agency's ADP facility is to provide the computing resources needed to fulfill the agency's mission requirements in the most cost-effective manner. Therefore, the workload forecasting process should also be performed in view of the benefit to the organization as a whole. The result of the workload forecasting process should provide the input for determining these computing resource needs.

Usually, Level 2 is the highest level to be considered by the workload forecasting team. However, the information that is available from Level 1 is useful in the analysis of the forecast results, especially for long-range forecasting. A method of translating an agency's mission requirements into computing resource requirements is described in the following sections.

Usually, several functions (Level 2) need to be performed by an agency in order to fulfill its mission. For illustration purposes, the payroll function is used to show how the translation can take place (see Figure 8).

The methods used for workload forecasting through the different levels are largely determined by the data availability and the relationship that can be established among the workload levels. In general, the workload forecasting can be performed by either using these relationships (if they are known or can be determined) through causal models or forecasting the workload independently on a given workload level. With causal models the effects (dependent variable) can be forecasted on the basis of their causes (independent variable(s)). In general, the causes can be forecasted in a relatively more reliable manner than the effects. However, in either case, the workloads need to be forecasted on the individual levels, either by using relationships that can be identified on a given level among the variables or forecasting a single variable as the function of time, using historical data and assuming that the past trend will continue into the future. The workload forecast can usually be performed on three different levels of the workload:

Level 3. quantifiable events;

Level 4. ADP operations performed;

Level 5. resources consumed.

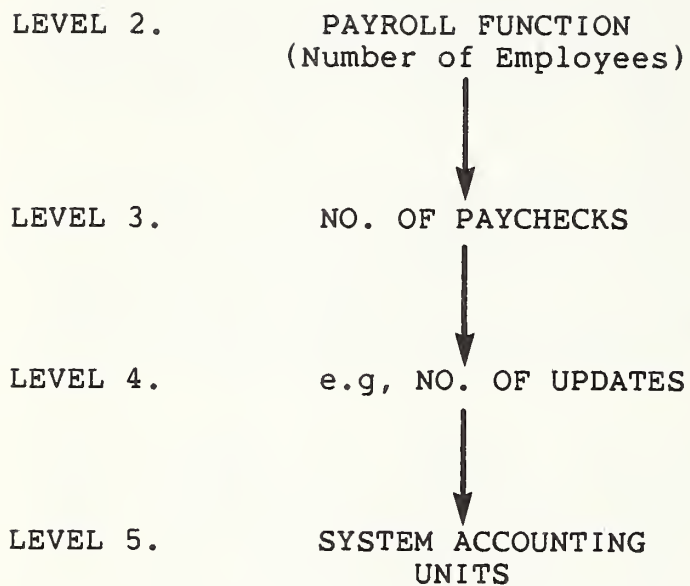


FIGURE 8. PAYROLL DESCRIPTION FOR THE
FORECASTING PROCESS

3.1.1 Forecast Level 3 Workload Requirements

As noted earlier, the workload on Level 3 can be described by the quantifiable events. As an example, the number of paychecks are the quantifiable events for the payroll function. They can be forecasted by using the relationships between Levels 2 and 3. For example, a simple linear regression model can be developed using historical data to determine these relationships. In this example the number of paychecks are the dependent variables and, for example, the personnel budget ceiling is the independent variable. If more than one independent variable can be identified on Level 2, and relationships can be established between these variables (independent) and the quantifiable events as the dependent variables, multiple linear regression as a causal model can be developed and used to forecast the quantifiable events. If no relationships can be identified, trend analysis techniques might be used to forecast the quantifiable events as a function of time.

3.1.2 Forecast Level 4 Workload Requirements

The workload on Level 4 can be described by the ADP operations necessary to produce a given quantifiable event. For example, the generation of a single paycheck might require one or more sorts, file updates, etc. On this level the workload requirements can be forecasted as on Level 3, either by causal models or by forecasting a given ADP operation as the function of time using trend analysis techniques.

3.1.3 Forecast Level 5 Workload Requirements

A computer system can be considered as a collection of resources upon which the workload places demands. The demands on these resources can be used to describe the workload on this level. This is the level at which the most quantifiable data exists. The workload requirements on this level can be forecasted in three (3) different ways:

1. by causal models, using data from Level 3 (as independent variables) and Level 5 data as the dependent variable (forecast unit);
2. by causal models, using data from Level 4 (as independent variables) and the Level 5 data as the dependent variable (forecast unit);
3. by trend analysis techniques, with the forecasting being performed on a single variable (on Level 5) as

the function of time using only historical data.

3.2 Apply Forecasting Techniques

In this section a practical example will be given on how to use each of the forecasting techniques discussed in the "Overview of Workload Forecasting Techniques" section. Table 1 can be used as a guide to determine the applicability of a given forecasting technique for the desired situation.

There are several commercially available forecasting software packages [DA83] to perform the necessary statistical computations.

3.2.1 Forecasting with Simple Moving Averages

The simple moving averages are most applicable for forecasting workload requirements during the operational phase of the ADP life-cycle. For illustration purposes, assume that an agency has twenty observations available for CPU utilization on a weekly basis. Table 2 is a list of the historical data for CPU utilization over the 20 weekly periods. For short term scheduling purposes, the operations personnel might be interested in what the CPU utilization would be in the 21st week.

As noted earlier, this method of forecasting can handle only stationary data. The stationary nature of the data can be determined by plotting the CPU utilization against the time periods. Figure 9 depicts that the CPU utilization for this example is nearly stationary.

After the stationary nature of the data has been determined, the decision has to be made concerning the number of observations (periods) to be included in the computation of the moving averages. The moving averages can be computed using formula:

$$F_t = \frac{Y_{t-1} + Y_{t-2} + \dots + Y_{t-n}}{n}$$

Where:

F_t = forecast value for time t ,
 n = number of observations included in the moving average,
 Y_t = actual (observation) value at time t .

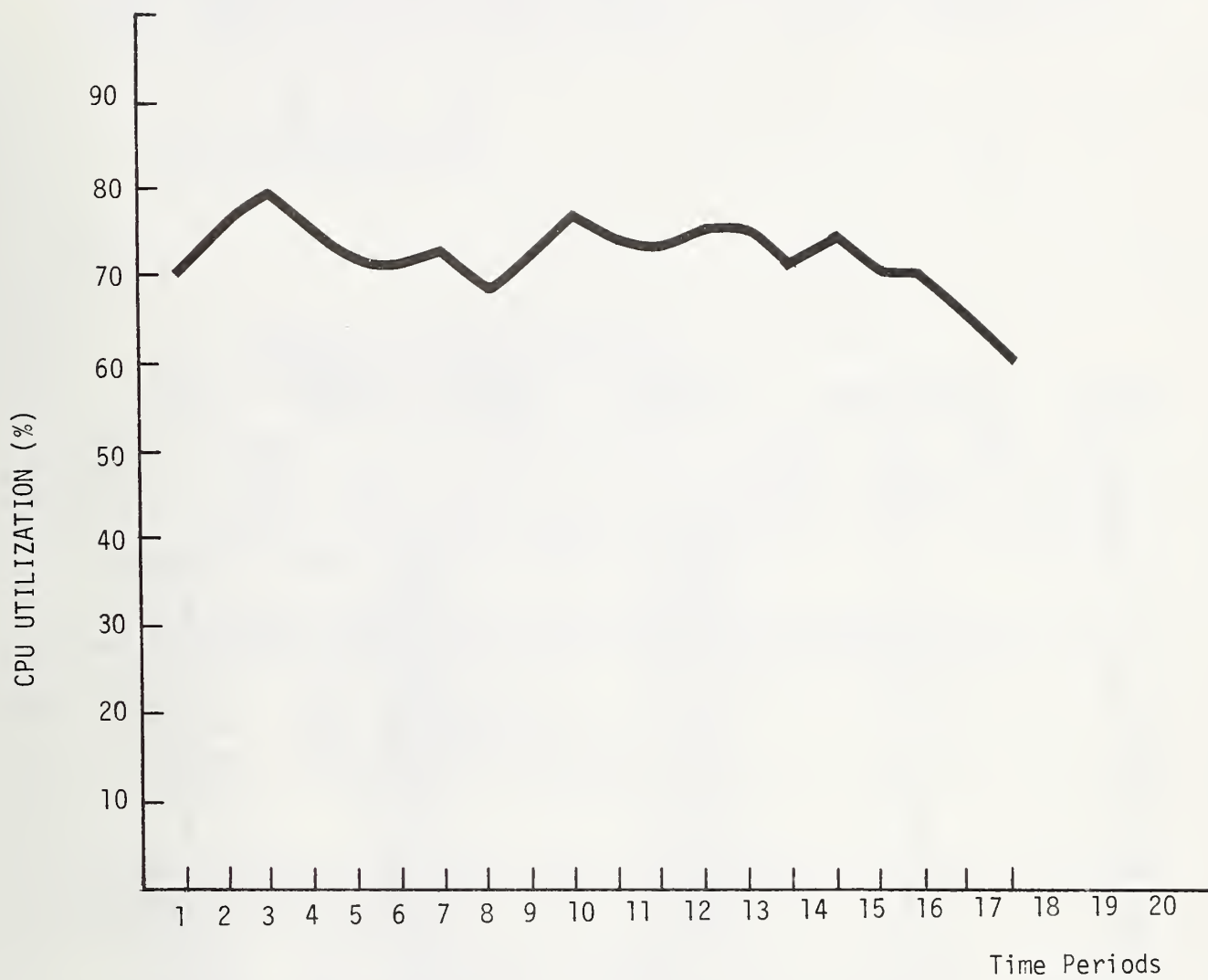


FIGURE 9. WEEKLY CPU UTILIZATION

PERIOD	OBSERVATION	FORECAST	FORECAST ERROR	% ERROR
1	71			
2	76			
3	79			
4	75	75.33	-0.33	- 0.33
5	72	76.67	-4.67	- 6.48
6	72	75.33	-3.33	- 4.62
7	73	73.00	0.00	0.00
8	69	72.33	-3.23	- 4.68
9	73	71.33	+0.67	+ 2.28
10	77	71.67	+5.33	+ 6.92
11	74	73.00	+1.00	+ 1.35
12	73	74.67	-1.67	- 2.29
13	75	74.67	+0.33	+ 0.44
14	75	74.00	+1.00	+ 1.33
15	71	74.33	-3.33	- 4.69
16	74	73.67	+0.33	+ 0.45
17	70	73.33	-3.33	- 4.76
18	70	71.67	-1.67	- 2.39
19	65	71.33	-6.33	- 9.74
20	60	68.33	-8.33	-13.88

TABLE 2. THREE MONTHS MOVING AVERAGE RESULTS
FOR CPU UTILIZATION

The number of periods to be included in the computation of the moving averages can be determined by varying the number of periods and choosing the one which gives the smallest mean squared error on the forecast. The mean squared error (MSE) can be computed using formula:

$$MSE = \frac{\sum_{t=1}^N (Y_t - F_t)^2}{N}$$

Where:

Y_t = actual (observations) value at time t ,
 F_t = forecast value at time t ,
 N = number of forecast values.

In this example, three observations (a test was made with three, five, and seven) in the moving average gives the smallest mean square error on the forecast. Table 2 shows the results of using three observations in the moving average to compute the forecast values.

In order to estimate the error on forecast values the mean absolute percentage error (MAPE) can be computed:

$$MAPE = \frac{\sum_{t=1}^N \left| \frac{(Y_t - F_t)}{Y_t} \right|}{N} (100).$$

In this example the mean absolute percentage error on the forecast is 3.93%.

Also, the forecaster might want to know if any consistent under- or over-estimation might occur. That is if a bias can be expected in the forecast results. This can be estimated by the mean percentage error (MPE) using formula:

$$MPE = \frac{\sum_{t=1}^N \left(\frac{Y_t - F_t}{Y_t} \right)}{N} (100).$$

The mean percentage error in this example is: -2.42%. That is a -2.42% underestimation can be expected in the forecast results.

It should be noted that MSE, MPE, and MAPE can also be used with other forecasting techniques.

The next step is to perform the forecast for period 21, using formula:

$$F_{t+1} = \frac{Y_t}{n} - \frac{Y_{t-n}}{n} + F_t ,$$
$$F_{21} = \frac{60}{3} - \frac{70}{3} + 68.33 ,$$
$$F_{21} = 65.00 .$$

Where:

n = number of observations included in the moving average,
 F_t = forecast for time t,
 Y_t = actual (observed) value at time t.

The forecast value of CPU utilization for time period 21 is 65%.

The major limitation of this technique is that a new forecast cannot be made until the forecast value for time period $t+1$ becomes a known value (actual observation).

3.2.2 Forecasting with Exponential Smoothing

One of the major problems with this method of forecasting is the selection of the appropriate weight. The smoothing weight is determined to be between zero and one. There is no best method for choosing the proper smoothing weight. In most cases, this is done experimentally by using different weights to decide which one is the most appropriate.

The exponential smoothing technique, like simple moving averages, is most appropriate for forecasting workload requirements during the operational phase of the ADP life-cycle. The method of performing workload forecasting with exponential smoothing is similar to the simple moving averages method. However, with this method of forecasting the more recent observations are given exponentially increasing weight [KE76].

As with moving averages, the time-series data (observations) need to be plotted in order to determine the stationary nature of the series. The next step is to determine the smoothing weight which gives the smallest mean squared error (MSE) on the forecast (See section 3.2.1). This can be accomplished by trial and error, giving

different values to the smoothing weights and choosing the one which gives the smallest mean squared error on the forecast. The determination of the most appropriate smoothing weight can be performed on the historical data, where the forecast values can be compared to the actual (historical) values. Table 3 shows the results of using 0.2, 0.5, and 0.7 as the smoothing weights in the exponential smoothing to compute the forecast values. Once the smoothing weight is determined, this method of forecasting requires only the latest observed value and the forecast for this same time period. The following equation can be used to perform the forecast: [MA78].

$$S_{t+1} = S_t + \alpha(X_t - S_t) .$$

Where:

F_t = forecast for time t,
 Y_t = actual (observed) value at time t,
 α = weighting (smoothing) value.

It is found that, using 0.7 as the smoothing weight gives the smallest mean squared error among the three smoothing weights tried. The mean absolute percentage error on the forecast is 3.91%, using 0.7 as the smoothing weight.

Now, the forecast can be made for time period 21. The forecast value of the CPU utilization for time period 21 is 60.90%.

3.2.3 Forecasting with Classical Decomposition Method

For the purpose of forecasting, it is assumed that the historical data can be described as:

$$S = T * C * I .$$

Where:

S = forecast value,
 T = trend factor,
 C = cyclical factor,
 I = seasonal factor,
 $*$ = denotes multiplication.

In the real world, randomness always exists in time series data. However, randomness cannot be predicted. Therefore, it is not included in the above equation. In order to use this equation in forecasting workload requirements the components (T , C , and I) must be

PERIOD	OBSERVATION	F O R E C A S T V A L U E		
		$\alpha = 0.2$	$\alpha = 0.5$	$\alpha = 0.7$
1	71			
2	76	71.0	71.0	71.0
3	79	72.0	73.5	73.4
4	75	73.4	74.3	77.3
5	72	73.7	73.1	75.7
6	72	73.4	76.6	73.1
7	73	73.1	72.8	72.3
8	69	73.1	72.9	72.8
9	73	72.3	71.0	70.1
10	77	72.4	72.0	72.1
11	74	73.3	74.5	75.5
12	73	73.4	74.2	74.4
13	75	73.3	73.6	73.5
14	75	73.6	74.3	74.5
15	71	73.9	74.6	74.8
16	74	73.3	72.8	72.1
17	70	73.4	73.4	73.4
18	70	72.7	71.7	71.0
19	65	72.2	70.8	70.3
20	60	70.7	67.9	63.1

TABLE 3. FORECAST RESULTS BY EXPONENTIAL SMOOTHING

determined. A detailed description for computing these components is given in the following paragraphs.

Table 4 shows monthly historical data over a five-year period for the number of transactions processed. The analysis of the historical data in Table 4 suggests that in this time series a seasonal pattern exists. This time series, if plotted, does not show a definite trend. The cyclical factor is not as apparent; its existence can usually be determined only after the decomposition of the time series is performed.

In the classical decomposition, first the seasonal factor is determined. This can be accomplished by:

1. Computing the 12-month moving average (MA), shown in Table 5. Each of these values are computed as the average of the preceeding six (6) historical data and the six (6) historical data which are following it.

2. As it can be seen in Table 5, the moving averages fall between each pair of months. It is necessary to adjust these moving averages so they will be in accord with the historical data. This can be accomplished through centering, by computing the 2-month moving average of the 12-month moving average. This is the 12-month centered moving average.

3. The next step is to compute the seasonal factor (Table 5). The seasonal factor is computed by expressing each historical data (Table 4) as a percentage of the corresponding 12-month centered moving average (Table 5). Then the seasonal index is computed for each of the 12 months by taking the average of the seasonal factors. The seasonal index is shown in Table 6.

The next step is to determine the trend factor in the historical data (Table 4). This can be accomplished by using a simple linear regression model (see 3.2.4.1 of Section III.). The trend equation is:

$$Y_t = 592.74 + 0.61X_t.$$

The trend values can then be computed for each month using this equation. For the simplicity of computation, in this example 593 is used as the trend value.

After the trend values have been computed, the cyclical factor can be computed next. The cyclical factor is calculated by dividing the centered 12-month moving average values by the trend values and the results are multiplied by 100. Next, the forecasting can be done for the desired point into the future using the identified pattern components. Assume the forecast is desired for January 1982. The seasonal index for January is 132.72. The trend

	1976	1977	1978	1979	1980
January	582	1022	822	724	790
February	478	780	624	610	612
March	468	716	542	536	538
April	490	766	586	578	610
May	482	694	520	542	546
June	434	688	514	520	550
July	536	716	566	618	662
August	488	636	488	474	492
September	482	620	436	490	602
October	528	634	566	472	628
November	702	652	532	552	750
December	812	740	650	696	1100

TABLE 4. MONTHLY DATA FOR THE NUMBER OF TRANSACTIONS

	12-MONTH MOVING AVERAGE	CENTERED 12-MONTH MOVING AVERAGE	SEASONAL FACTOR	CYCLICAL FACTOR
1976				
January				
February				
March				
April				
May				
June	540.17			
July	576.83	558.50	95.97	94.18
August	602.00	589.42	82.79	99.39
September	626.67	614.33	78.45	103.77
October	645.67	636.17	82.99	107.28
November	663.33	654.50	107.26	110.37
December		673.92	120.49	113.64
1977				
January	684.50	692.00	147.69	116.69
February	699.50	705.67	110.53	119.00
March	711.83	711.58	100.62	119.99
April	723.33	727.75	105.26	122.72
May	732.17	730.01	95.07	123.10
June	728.00	725.00	94.90	122.26
July	722.00	697.42	105.38	114.57
August	636.83	664.58	98.67	112.07
September	692.33	685.08	90.50	115.53
October	677.83	670.33	94.58	113.04
November	662.83	655.58	99.45	110.55
December	648.33	666.08	111.10	112.32

TABLE 5. MOVING AVERAGES AND CYCLICAL FACTORS FOR THE
NUMBER OF TRANSACTIONS PROCESSED

	12-MONTH MOVING AVERAGE	CENTERED 12-MONTH MOVING AVERAGE	SEASONAL FACTOR	CYCLICAL FACTOR
1978				
	683.83			
January		652.58	125.96	110.04
	621.33			
February		615.17	101.44	103.74
	609.00			
March		603.34	89.83	101.74
	597.67			
April		592.84	98.85	99.97
	588.00			
May		583.00	89.19	98.31
	578.00			
June		574.25	89.51	96.84
	570.50			
July		566.42	99.93	95.52
	562.33			
August		561.75	86.87	94.73
	561.17			
September		560.92	77.73	94.59
	560.67			
October		560.33	101.01	94.49
	560.00			
November		562.92	94.56	94.93
	561.83			
December		562.08	115.64	94.79
1979				
	562.33			
January		564.50	128.26	95.19
	566.67			
February		566.09	107.26	95.46
	565.50			
March		567.75	94.41	95.74
	570.00			
April		566.09	102.10	95.46
	562.17			
May		563.00	96.27	94.94
	563.83			

TABLE 5. MOVING AVERAGES AND CYCLICAL FACTORS FOR THE
NUMBER OF TRANSACTIONS PROCESSED

	12-MONTH MOVING AVERAGE	CENTERED 12-MONTH MOVING AVERAGE	SEASONAL FACTOR	CYCLICAL FACTOR
1979				
June	567.67	565.75	91.75	95.40
July	573.17	570.42	108.34	96.19
August	573.17	573.17	82.70	96.66
September	573.50	573.34	85.46	96.68
October	576.17	574.84	82.11	96.94
November	576.50	576.34	95.78	97.19
December		577.75	120.47	97.43
1980				
January	579.00	580.84	136.01	97.95
February	582.67	583.42	104.90	98.38
March	584.17	568.84	94.58	95.92
April	553.50	580.00	105.17	97.81
May	606.50	614.75	88.82	103.67
June	623.00	639.84	85.96	107.90
	656.67			

TABLE 5. MOVING AVERAGES AND CYCLICAL FACTORS FOR THE
NUMBER OF TRANSACTIONS PROCESSED

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.
1976							95.97	82.79	78.45	82.99	107.26	120.49
1977	147.69	110.53	100.62	105.26	95.07	94.90	105.38	98.67	90.50	94.58	99.45	111.10
1978	125.96	101.44	89.83	98.85	89.19	89.51	99.93	86.67	77.63	101.02	94.56	115.64
1979	128.26	107.76	94.41	102.10	96.27	91.75	108.34	82.70	85.46	82.11	95.78	120.47
1980	136.01	104.90	94.58	105.17	88.82	85.96						
Average ⁽¹⁾	132.14	106.33	94.50	105.22	92.13	90.63	102.66	84.73	81.96	88.79	97.62	118.06 1194.77
Adjusted ⁽²⁾												
Seasonal Index	132.72	106.80	94.92	105.69	92.54	91.03	103.11	85.10	82.32	39.18	98.05	118.58 1200.04

(1) The average excluding the maximum and minimum values

(2) seasonal adjustment factor: 1200.04/1194.77

TABLE 6. SEASONAL FACTORS FOR THE NUMBER OF TRANSACTIONS PROCESSED

value is approximately 593. The cyclical factor needs to be estimated by the forecaster. The forecaster should examine expected changes in the workload such as: introduction of new application systems, elimination of application systems and any other factors that might effect the cyclical component in the forecasting process. The cyclical factor in this example is fairly smooth, does not exhibit any radical change over a short period of time. In this example a somewhat subjectively higher cyclical factor is used (1.09) than the last computed cyclical factor, 107.90 (for June 1980).

The model to be used in forecasting is:

$$S = 593 * 1.32 * 1.09.$$

The forecast value for January 1982 is: 853.21.

3.2.4 Forecasting with Simple Linear Regression

The simple regression method can be used for forecasting workload requirements throughout the life-cycle of an ADP system. The simple regression analysis technique can be used either as a trend analysis tool or as a causal model. When it is used as a trend analysis technique, the values of the independent variable are the time periods, and the dependent variable is the unit of measure for which the trend is to be determined. In the case of a causal model, the independent variable in the regression equation is a variable which causes a linear rate of change in the dependent variable (forecast unit of measure).

3.2.4.1 Simple Linear Regression as a Causal Model

As noted earlier, simple regression analysis is a useful tool for quantifying the relationship between two variables. Correlation analysis can be used to determine the linearity of this relationship. This relationship can be used for forecasting the dependent variable. When using simple linear regression as a causal model, a linear relationship is assumed between the dependent (Y) variable and the independent (X) variable). Table 7 contains 30 observations for the number of transactions processed (as dependent variable), and the number of orders received (as the independent variable) .

The most commonly used regression method to find the best fit line to the points is the least squares fit technique. The simple linear regression equation between

PERIOD	NO. OF TRANSACTIONS	NO. OF ORDERS
1	46	20
2	48	14
3	66	14
4	128	38
5	130	36
6	154	32
7	164	48
8	166	50
9	172	56
10	176	52
11	198	42
12	210	48
13	302	80
14	324	78
15	754	186
16	390	100
17	398	96
18	412	116
19	420	104
20	438	110
21	442	120
22	448	120
23	508	140
24	512	146
25	562	148
26	576	136
27	694	192
28	738	188
29	754	186
30	784	192

TABLE 7. HISTORICAL DATA OVER 30 PERIODS

two variables Y (dependent variable) and X (independent variable) is denoted:

$$Y = a + bX.$$

Where "a" is the Y intercept (the value of Y when $X = 0$), "b" is the regression coefficient (or the slope of the line) which indicates the change in Y (dependent variable) when X (independent variable) changes by one unit.

This procedure is simple and can be dealt with without computers. Several advanced hand calculators have the capability to perform these statistical computations.

The parameters "a" and "b" can be determined by the least squares fit method. With the least squares fit method the distance between the observations and the corresponding points on the straight line is minimized. The parameters can be obtained by the following equations:

$$b = \frac{\sum_{i=1}^N x_i y_i - \bar{X} \sum_{i=1}^N y_i}{\sum_{i=1}^N x_i^2 - \bar{X} \sum_{i=1}^N x_i} ; \quad a = \bar{Y} - b\bar{X} .$$

Using these equations and the historical data from Table 7, the following values are obtained for the parameters:

$$\begin{aligned} a &= -7.81, \\ b &= 3.93. \end{aligned}$$

Where:

N = number of observations,
 \bar{X} = mean value of Xs,
 \bar{Y} = mean value of Ys.

The causal model (regression equation) derived by the least squares fit method is:

$$F_i = -7.81 + 3.93X_i$$

Where:

F_i = the computed (fitted) values from the regression equation.

Before the regression equation can be used for forecasting the number of transactions processed, correlation analysis needs to be performed to determine the degree of association between the two variables (dependent and independent). The linear association between two variables can also be visually detected by using a scatter diagram, where each pair of observations (dependent and independent variable) is plotted as a point. A straight line could be manually fitted to these points, but this line would not necessarily be the best fitted one.

This association can be measured by the correlation coefficient of the regression line. The correlation coefficient ranges between -1 and +1. The closer the correlation coefficient is to +1 (for our purpose), the better the association is between the two variables. It is important to note that this measure is not meaningful unless sufficient historical data is available.

The equation that can be used for computing the correlation coefficient is:

$$r = \frac{\sum_{i=1}^N (F_i - \bar{Y})^2}{\sum_{i=1}^N (Y_i - \bar{Y})^2}$$

Where:

- Y_i = the observed values (dependent variables),
- F_i = the computed (fitted) values from the regression equation,
- \bar{Y} = mean value of Y_s ,
- N = number of observations.

The correlation coefficient is: 0.99. Since a large number of observations (30) was used to determine the correlation coefficient, it is reasonable to assume that the high correlation (0.99) is significant in this example.

Also, a statistical significance test can be performed on the regression coefficient. The significance of the regression coefficient can be determined by its standard error. It can be computed using [MA78]:

$$SE_b = \frac{\sum_{i=1}^N (Y_i - F_i)^2 / N / 2}{\sum_{i=1}^N (X_i - \bar{X})^2} = \sqrt{\frac{879}{97671}} = \frac{29}{312} = 0.09$$

The standard error of the regression coefficient (b) is 0.09.

In order to obtain confidence in the regression coefficient and for the predicted values, statistical significance tests need to be performed. After the standard error of the regression coefficient is computed, a t-test should be performed to determine whether the regression coefficient differs significantly from zero. The t-test for "b" can be computed as follows:

$$t\text{-test}_b = \frac{b}{SE_b} = \frac{3.93}{0.09} = 43 .$$

By comparing the computed value of t to the appropriate value in a t-table, the significance level can be determined for the regression coefficient. Since the t-table value [SP75] is greater than two and the computed t-test for "b" is greater than the critical t-table value, it can be concluded (with almost 100% certainty) that the regression coefficient is significantly different from zero.

In addition, the "goodness of fit" should also be determined. This can be measured by the coefficient of determination (square of the correlation coefficient). The coefficient of determination measures the proportion of total variation about the mean Y that is accounted for in the regression equation (explained variation). This can be computed as:

$$R^2 = \frac{\sum_{i=1}^N (F_i - \bar{Y})^2}{\sum_{i=1}^N (Y_i - \bar{Y})^2}$$

that is, the sum of the squares of the deviations of the computed (estimated) values from their mean.

In this example, the coefficient of determination is 0.98. That is 98% of the variation from the mean value of Y is explained by the model. The significance (at 95% confidence level) of this explained variation can be obtained by the F-test:

$$F = \frac{\sum_{i=1}^N (F_i - \bar{Y})^2 (K-1)}{\sum_{i=1}^N (Y_i - F_i)^2 / (N-K)}$$

Where:

N = number of observations,
K = number of variables.

This value can be compared to the appropriate entry (intersection of 1 and 28 in the F-table) [SP75] which is 4.20. Since the computed F-test is greater than five (as a general rule) [MA78], and it is greater than the critical value obtained from the F-table, the forecaster can be 95% certain that the regression equation explains 98% of the total variation from the mean value of Y.

The unexplained variation is the sum of the squares of the deviations of the Y values from their computed (estimated) values:

$$\sum_{i=1}^N (Y_i - F_i)^2.$$

These unexplained variations are called the residuals. An examination of the residuals (difference between observed and predicted values) also needs to be performed. At least three criteria need to be satisfied:

1. the mean value of residuals must be zero;
2. the residuals must be normally distributed; and
3. the variance of residuals $\sum_{i=1}^N \left(\frac{F_i - \bar{Y}}{N-1} \right)^2$ must be constant.

In most cases, the forecaster and the decision makers are interested in the confidence interval (bounds) on the forecast. This can be obtained by computing the standard error on the forecast. The standard error on the forecast can be computed by using formula:

$$SE_f = \left[\frac{\sum_{i=1}^N (Y_i - F_i)^2}{N-2} \right] \left[1 + \frac{1}{N} + \frac{(X_i - \bar{X})^2}{\sum_{i=1}^N (X_i - \bar{X})^2} \right]$$

$$SE_f = \left(\frac{24628}{28} \right) \left[1 + \frac{1}{30} + \frac{(50 - 96.27)^2}{97671} \right] = 4.16 .$$

When 50 orders are received, the number of transactions processed (at 95% confidence level) would be:

$$Y_f = -7.81 + 3.93(50) \pm (4.16)$$

bound on the forecast.

That is, the number of transactions processed for 50 orders received would be between 180 and 197 with 95% certainty.

3.2.4.2 Trend Analysis with Simple Linear Regression

The method of using simple regression for trend analysis is similar to the use of simple regression as a causal model. The only exception is that, with this method, the values of the independent variable are the time periods.

3.2.5 Forecasting with Multiple Linear Regression

The multiple regression method, like simple regression, can be used for forecasting workload requirements throughout the life-cycle of an ADP system. The forecasting with multiple regression in basic concept is similar to the method of forecasting with simple regression. However, in the case of multiple regression, the dependent variable (forecast unit of measure) is defined as a function of several independent variables. The independent variables are assumed to have linear relationships with the dependent variable.

In multiple linear regression, the theoretical relationship between the dependent and independent variables is assumed to be of the form:

$$Y = a + b_1 x_1 + b_2 x_2 + \dots + b_i x_i + \alpha$$

Where, Y is the dependent variable, x_1, x_2, \dots, x_p are the independent variables, and α is a normally distributed random error (residual) with zero mean and constant variance. For the relationship to be modeled in workload forecasting, in most cases it is reasonable to assume that the relationships may be of this form. The assumption of the linear relationships allows the building of a model for making predictions. Using historical data (observations) for the variables by the least squares fit method, the coefficients for the independent variables can be estimated. The residual can be computed as the difference between the estimated value of Y and the observed value of Y.

The choice of the independent variables should be based on two criteria:

1. expected linear relationship with the dependent variable; and
2. availability of estimates for the future.

Although non-linear relationships can be transformed into linear, the model might not be as successful in forecasting as it would be if linear relationships existed. Simple techniques, such as scatter-diagrams, can be used to initially determine the linearity.

For the relationships to be modeled in the area of workload forecasting it is reasonable to assume, in most cases, that the relationships may be in this form. For example, the processing requirements to support a payroll system may reasonably be expected to be a linear function of the number of people the system supports.

It should be noted that the forecaster planning to employ this technique should obtain a statistical package to perform the necessary computations.

In order to develop a multiple regression model for the purpose of forecasting workload requirements, the forecast unit of measure and the candidate independent variables need to be identified.

The selection of the variables can be performed by stepwise regression [DR66]. This procedure enters the independent variables into the regression equation one at a time. At each step those variables are entered which are making the largest contribution among those not yet entered. With the introduction of the next variable, the contribution of some of the previously introduced variables might become negligible. That is, they can be eliminated from the equation. Hence, after each step, the contribution of the independent variables are reevaluted.

Usually, in practice, the procedure begins by computing the simple correlation of each independent variable with the dependent variable and then first introducing the independent variable with the highest correlation into the regression equation. Those variables which have the largest partial F-statistic (F-test) are then added one at a time. The procedure terminates when the F-statistic of the variables not yet introduced are statistically not significant.

The example given below uses the results of a previous study [PO78]. The dependent variable to be modeled is the total number of I/O's, and the candidate independent

variables are shown in Table 8.

In modeling the total number of I/O's, stepwise regression is performed with all the independent variables shown in Table 8. In Table 8 it can be seen that Step 1 selects Airmen as the best single independent variable. The coefficient of determination for this regression is 0.53; the standard error is 4.1×10^6 ; the mean value of total I/O's is 22.7×10^6 . Adding the Account Control in the second step improves the "goodness of fit" considerably, the coefficient of determination increases to 0.76 and decreases the standard error to 2.93×10^6 . The procedure terminates at this point if the inclusion of the variables is restricted to those variables whose coefficients are significant at the 0.01 level, as measured by the partial F-statistic. Allowing a slightly less significant variable, the procedure next enters Civil Engineering, which increases the coefficient of determination to 0.78 and lowers the standard error to 2.82×10^6 . The coefficient of the added variable (Civil Engineering) has a partial F-statistic of 6.2, significant at 0.02 level. The fourth step adds Medical Material, bringing the determination of coefficient to just under 0.80 and reducing the standard error of the previous equation by two percent. The least significant coefficient, of the Medical Material variable just entered, has a partial F-statistic of 4.1, significant at the 0.05 level.

The procedure would stop at this point if the level for entry were set at the 0.05 level. If less significant terms are allowed to enter, steps 5, 6, and 7 add Data Control, Fighter Pilots, and Travel. Each brings a slight increase in the coefficient of determination and decrease in the standard error. Steps 8 and 9 then remove the Medical Material and Data Control variables, leaving five variables in the equation, achieving 0.82 for the coefficient of determination and standard error of 2.65×10^6 . Each variable of this equation is significant at the 0.01 level.

As can be seen, the stepwise regression procedure is a useful tool for selecting the most appropriate model. The equation in step 4 provides the best model. The partial F-statistic of the coefficients of the variables are all significant at the 0.05 level, all but one being significant at the 0.01 level.

In practice, the selection of the most appropriate model is not always straightforward. The examination of all possible equations does not provide an incisive answer. It should be remembered that the independent variables included in the equation to forecast the values for the dependent variable are themselves the result of a separate forecasting or experts' opinions. That is, when selecting the most appropriate model, the degree of uncertainties associated with the independent variables need to be taken into account in addition to the statistical significance tests performed.

STEP	PARTIAL F-STATISTIC OF INDEPENDENT VARIABLES ENTERED (REMOVED)							DEGREES OF FREEDOM (N-K)	NO. OF VARIABLES	R ²	SE:10 ⁶
	AIRMEN	ACCOUNTS CONTROL	CIVIL ENGINEERING	MEDICAL MATERIAL	DATA CONTROL	FIGHTER PILOTS	TRAVEL				
1	78.7							70	1	0.529	4.11
2	164.9	69.0						69	2	0.765	2.93
3	73.4	58.0	6.2					68	3	0.784	2.82
4	71.0	50.0	7.2	4.1				67	4	0.797	2.76
5	20.8	34.8	9.2	4.6	3.9			66	5	0.808	2.70
6	11.9	36.7	12.1	3.2	4.6	3.2		65	6	0.817	2.66
7	10.8	11.8	12.3	2.7	1.6	5.1	3.6	64	7	0.827	2.61
8	24.1	11.7	11.0	2.4	(1.6)	5.3	6.6	65	6	0.822	2.62
9	23.2	12.1	10.8	(2.4)		7.1	7.2	66	5	0.816	2.65

TABLE 8. STEPWISE REGRESSION FOR TOTAL NUMBER OF I/O's

3.2.6 Forecasting with the Box-Jenkins Method

The procedure for the Box-Jenkins method in general can best be described by a schematic diagram, as depicted in Figure 10.

As noted earlier, and depicted by Figure 10, this method of forecasting can handle any data pattern. The identification of the pattern in the data is accomplished by the method itself. The pattern identification leads to the model to be employed. The primary aid in the identification of the model to be utilized is the autocorrelation. High autocorrelation indicates seasonal or cyclical pattern, and where the autocorrelation is low or does not exist, it indicates random data. The Box-Jenkins method uses this information to arrive at the optimal model to be used in forecasting. Further information on the use of autocorrelation can be found in [B076]. The autocorrelation can be computed by the following formula [KE76]:

$$r_k = \frac{\frac{1}{n-k} \sum_{t=1}^k (x_t - \bar{x})(x_{t+k} - \bar{x})}{\frac{1}{n} \sum_{t=1}^k (x_t - \bar{x})^2}$$

Where:

n = the number of observations,
 X_t = the value of the variable at time t ,
 k = the length of time,
 \bar{X} = the mean value of all of the observations.

Three general classes of models can be utilized in the Box-Jenkins method of forecasting:

1. Autoregressive (AR),
2. Moving Average (MA), and
3. Autoregressive-Moving Average (ARMA).

The following is a brief description of these general classes of models.

The Autoregressive Model (AR(p)).

The autoregressive model can be characterized by the autocorrelation function which decays exponentially to zero or exhibits an exponentially damped sine wave. The

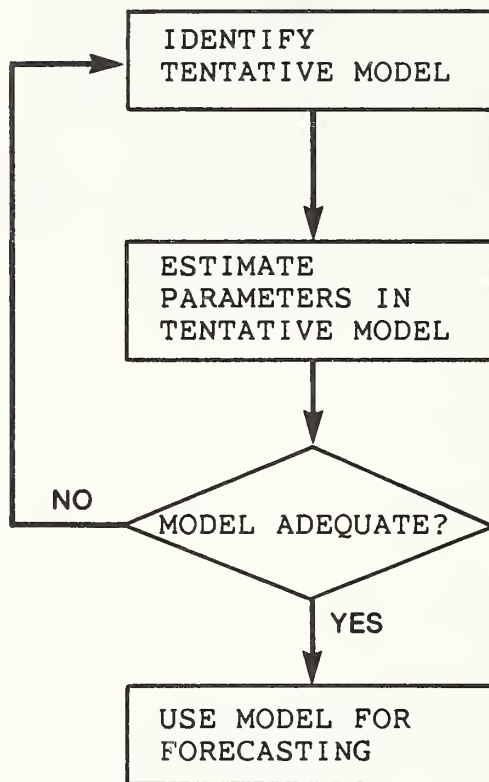


FIGURE 10. BOX-JENKINS METHOD OF FORECASTING

autocorrelation coefficient versus time period can be plotted on a correlogram to observe this behavior. A graph of the autocorrelation function is referred to as the correlogram. An autoregressive model of order p describes the current observation as a weighted sum of " p " previous observations plus a random perturbation or residual that cannot be explained by the model. The p order autoregressive model is in the form of:

$$X_t = \sum_{j=1}^p \Phi_j X_{t-j} + Z_t$$

Where:

X_t = observations
 Φ_j = the weighting parameter,
 Z_t = the residual.

The procedure for computing the parameters can be found in [NE73].

The Moving Average Model (MA(q))

A moving average model of order q can be characterized by the autocorrelation function which if plotted on a correlogram exhibits an abrupt cut off after the time period q . In the moving average model, the current value (dependent variable) is computed as the sum of the current residual plus the weighted sum of previous residuals. Mathematically, this can be expressed as:

$$X_t = Z_t + \sum_{k=1}^q \alpha_k Z_{t-k}$$

Where:

X_t = the value of the variable at time t ,
 α_k = the weighting parameter,
 Z_t = the residual,
 q = the number of previous residual in the weighted sum.

The Autoregressive-Moving Average Model (ARMA(p,q))

Sometimes it might be desirable to combine the autoregressive AR(p) and the moving average MA(q) models to describe a series of observations. Such a model is referred

to as autoregressive-moving average model. This model when used in forecasting bases the forecast value both on past observations and on the residuals. This model can be mathematically expressed as:

$$X_t = \sum_{j=1}^p \Phi_j X_{t-j} + \sum_{k=1}^q C_k Z_{t-k}$$

Where:

Φ = the weighting parameter.

Since this model is a mix of the autoregressive and moving average models, it can be characterized by an autocorrelation function which decays exponentially or exhibits an exponentially damped sine wave and/or shows an abrupt cut off after p-q periods.

This method of forecasting is applicable for forecasting workload requirements in the operational phase of an ADP system life-cycle. The Box-Jenkins method of forecasting does not require prior assumptions about the underlying pattern in the series.

The first step in the Box-Jenkins method of forecasting is the identification of the tentative model. This can be accomplished by the analysis of the autocorrelations.

Table 9 is a list of autocorrelations for illustration purposes. The examination of the autocorrelations shows peaks near time periods (lags) seven and fourteen. This indicates the possibility of a seven-day cycle (seasonal pattern), which suggests that seasonal differencing is necessary. By taking the seasonal difference of period seven it can be seen (Table 10) that all of the sample autocorrelations are near zero, except for period seven. This is an indication that the model is in the form

$$X_t = X_{t-7} + Z_t - \Phi Z_{t-7},$$

which is a seasonal model. It should be noted that the identification of the model is not as simple as it might seem from the above example. Users attempting to apply this method would find the following references [BO76,KU78,NE73] useful. Also, there are several software packages available to perform the required statistical computations.

The next step is to determine the Φ parameter. This can be accomplished by minimizing the sum of squares on the residuals (Z_t). It is found to be 0.6144.

<u>PERIODS</u>	<u>AUTOCORRELATIONS</u>				
1-7	0.1579	0.0849	-0.0447	0.0259	0.1680
8-14	0.0361	-0.0258	-0.0819	0.0910	0.1999
15-21	0.0402	-0.0041	-0.0342	0.0834	0.1243
22-25	-0.0599	-0.0287	-0.0024	0.0178	0.0621

TABLE 9. ESTIMATED AUTOCORRELATIONS OF UNDIFFERENCED LOGGED DATA

<u>PERIODS</u>	<u>AUTOCORRELATIONS</u>				
1-7	0.0403	0.0497	-0.0103	0.0153	0.0192
8-14	0.0444	-0.0237	-0.0576	0.0349	0.0711
15-21	0.0454	0.0115	0.0027	0.0332	-0.0015
22-25	0.0385	-0.0026	-0.0147	-0.0232	0.0121

TABLE 10. ESTIMATED AUTOCORRELATIONS OF DIFFERENCED LOGGED DATA

The model to be used in forecasting is thus:

$$X_t = X_{t-7} + Z_t - 0.6144Z_{t-7} .$$

Before this model can be used in forecasting, diagnostic checks need to be performed to determine the adequacy of the model. This can be accomplished by the analysis of autocorrelations of the estimated residuals. If the residuals do not show any pattern and they all are nearly zero the model seems adequate.

Since the model is seasonal, one more diagnostic check needs to be performed on the residuals using a cumulative periodogram [BO76], Figure 11 [KU78]. By a careful inspection of the cumulative periodogram plot of the residuals it can be determined whether any periodic components exists in the series which are not explained by the model. Kolmogorov-Smirnoff goodness of fit [BO76] can be performed to determine whether there are any periodicities in the residuals. The cumulative periodogram (Figure 11) does not show any periodicity in the residuals. Thus it can be assumed that the model is adequate.

STEP 4. Analyze Forecast Results

The results of the forecasting process need to be carefully analyzed and evaluated before any decisions can be made based on the forecast results. The results of the forecasting process present the decision makers with information on whether adequate computing resources are available to perform the users workload. The actions to be taken might involve one or more of the following activities:

- tune the current system,
- upgrade system components,
- obtain outside resources,
- replace the current system.

In the analysis and evaluation step of the forecasting process several important activities should take place.

The purpose of the workload forecasting process is to provide information on future processing requirements in order to provide adequate computing resources to perform the user's workload and ultimately to fulfill the agency's mission. Therefore, it must be determined whether the applications/application systems included in the forecast process are the ones which are necessary to support the agency's mission. This can be accomplished by submitting

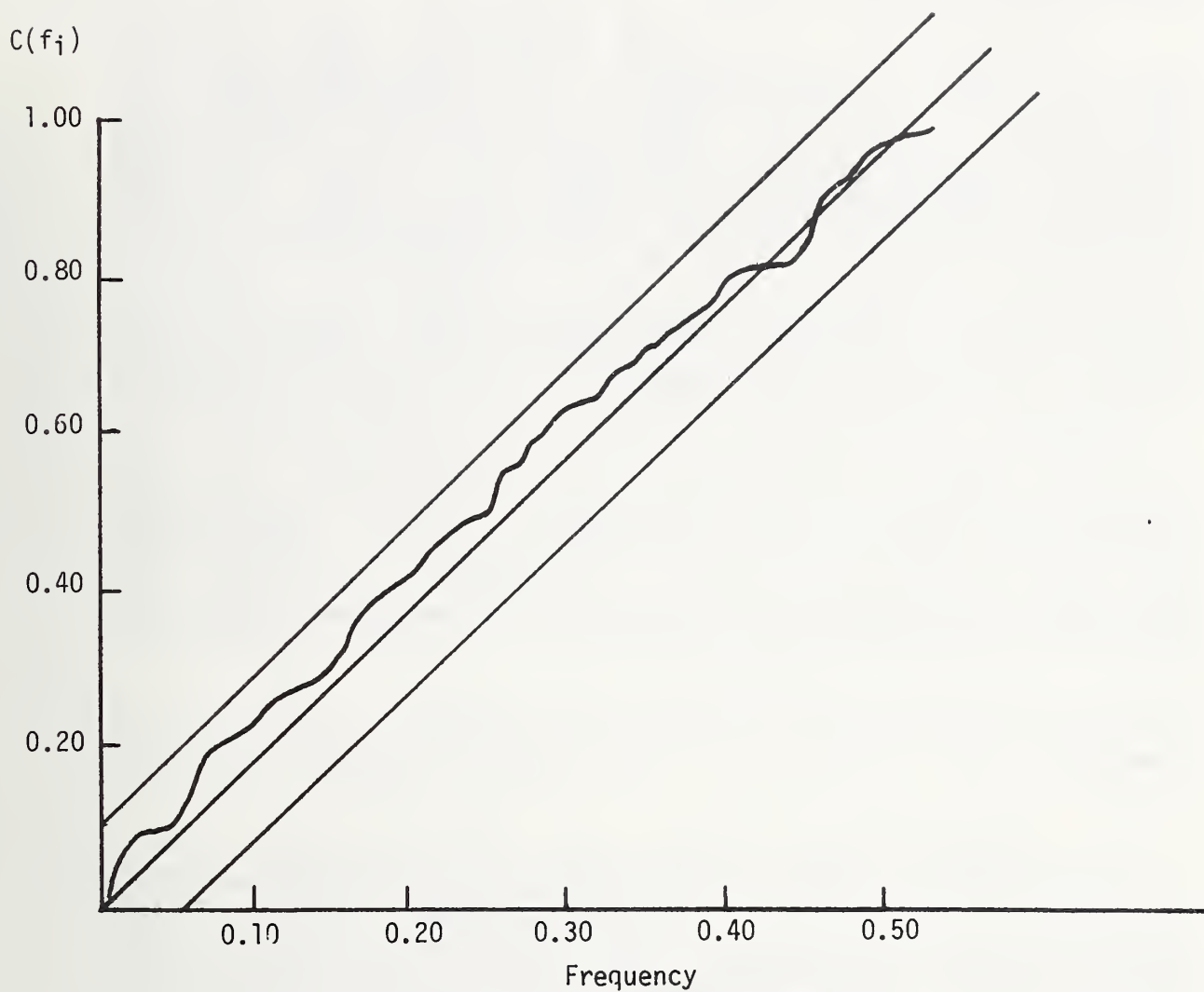


FIGURE 11. CUMULATIVE PERIODOGRAM OF RESIDUALS

the forecast results to personnel who have been surveyed in the data collection process.

Concerning the forecasting techniques used, several factors need to be analyzed. One of the most important considerations is the examination of the model used in the forecasting. That is, whether the conditions under which the model was derived are still the same. If not, the model needs to be updated. Also, if causal models (e.g., multiple regression) were used in the forecasting process, it is desirable to perform several forecasts by varying the values (sensitivity analysis) of the independent variables (predictors). This is usually desirable when the users have doubt about the certainty of their future requirements, and/or if the applications/application systems involved play a highly important role in fulfilling the agency's mission.

The forecast results should be presented to the decision makers by identifying the outcomes of the different techniques applied, the results of the alternative techniques, and sensitivity analysis, if one is performed. Also, the likely solutions (e.g., upgrade) to overcome any possible shortcomings in the computing resources, should be submitted to the decision makers.

SUMMARY

This guide attempted to provide practical guidance on workload forecasting in a step-by-step fashion. Some agencies may find that the sequence of these steps is not suited to their particular needs, in this case the Table of Contents can be used as a reference to the specific sections of interest. However these steps should at least be considered, especially by those agencies where no formal forecasting currently exists.

It is hoped that this guide will help agencies to obtain up-to-date information on their future workload requirements upon which informed management decisions can be made regarding computer resource needs to meet these future workload requirements.

APPENDIX A

SUGGESTED READINGS AND REFERENCES

- [BO76] Box, George E.P.; Jenkins, Gwilym M. "Time Series Analysis, Forecasting and Control." Revised Edition. Holden-Day, Inc. 1976.
- [CH74] Chambers, John C.; Mullick, Satinder K.; Smith, Donald D. "An Executive's Guide to Forecasting." John Wiley and Sons. 1974.
- [DA83] Datapro Research Corporation, "Management Sciences," Delran, NJ. 08075.
- [DR66] Draper, N. R. "Applied Regression Analysis." John Wiley and Sons. 1966.
- [KE76] Kendall, Maurice. "Time Series." Charles Griffin and Company LTD. 1976.
- [KU78] Kulp, Richard W.; Melendez, Kenneth. "An Application of Time Series Analysis in Computer Performance Evaluation." Computer Performance Evaluation Users Group: CPEUG 14th Meeting; 1978.
- [LI75] Linstone, Harold A.; Turoff, Murray, eds. "The Delphi Method: Techniques and Applications." Addison-Wesley, Reading, Ma. 1975.
- [MA78] Makridakis, Spyros; Wheelright, Steven C. "Interactive Forecasting Univariate and Multivariate Methods." Holden-Day, Inc. 1978.
- [NB80] National Bureau of Standards. Federal Information Processing Standards (FIPS) Publication. "Guideline on Constructing Benchmarks for ADP System Acquisitions." NBS. 1980 September 18; FIPS PUB 75.
- [NE73] Nelson, Charles R. "Applied Time Series Analysis for Managerial Forecasting." Holden-Day, Inc. 1973.
- [PO78] Poggio, Eugene C.; Ainslie, Naomi. "Forecasting Processing Requirements of Standard USAF Base Level Computer Systems." 1978.
- [SP75] Spiegel, Murray R. "Theory and Problems of Probability and Statistics." McGraw-Hill Book Company, 1975.
- [SA79] Sarna, David E.Y. "Forecasting Resource Utilization Using Key Volume Indicators." Proceedings National Computer Conference; 1979.

[SH67] Shiskin, Julius; Young, Allen H.; Musgrave, John C.
"The X-11 Variant of the Census Method II Seasonal
Adjustment Program." U.S. Department of Commerce, Bureau of
Census. Technical Paper No.15, 1967.

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET (See instructions)		1. PUBLICATION OR REPORT NO. NBS/SP-500/123	2. Performing Organ. Report No.	3. Publication Date March 1985
4. TITLE AND SUBTITLE Computer Science and Technology: Guide on Workload Forecasting				
5. AUTHOR(S) Helen Letmanyi				
6. PERFORMING ORGANIZATION (If joint or other than NBS, see instructions) NATIONAL BUREAU OF STANDARDS DEPARTMENT OF COMMERCE GAITHERSBURG, MD 20899			7. Contract/Grant No.	8. Type of Report & Period Covered Final
9. SPONSORING ORGANIZATION NAME AND COMPLETE ADDRESS (Street, City, State, ZIP) Same as in item 6 above.				
10. SUPPLEMENTARY NOTES Library of Congress Catalog Card Number: 85-600504 <input type="checkbox"/> Document describes a computer program; SF-185, FIPS Software Summary, is attached.				
11. ABSTRACT (A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here) The purpose of this guide is to provide ADP managers and technical personnel with useful quantitative techniques for forecasting future workload requirements. It additionally provides a step-by-step approach to the forecasting process. Readers can then, in a timely manner, provide the computing resources needed to perform the user's workload at required service levels throughout the life-cycle of an ADP system. These techniques are described so that readers with little or no training in statistics should find them useful. However, this guide does not intend to give an exhaustive treatment of the techniques discussed. Readers requiring more information are referred to Appendix A ("Suggested Readings and References").				
12. KEY WORDS (Six to twelve entries; alphabetical order; capitalize only proper names; and separate key words by semicolons) causal models; forecast spans; time-series models; workload forecasting techniques; workload levels; workload transition.				
13. AVAILABILITY <input checked="" type="checkbox"/> Unlimited <input type="checkbox"/> For Official Distribution. Do Not Release to NTIS <input checked="" type="checkbox"/> Order From Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. <input type="checkbox"/> Order From National Technical Information Service (NTIS), Springfield, VA. 22161			14. NO. OF PRINTED PAGES 68 15. Price	

**ANNOUNCEMENT OF NEW PUBLICATIONS ON
COMPUTER SCIENCE & TECHNOLOGY**

Superintendent of Documents,
Government Printing Office,
Washington, DC 20402

Dear Sir:

Please add my name to the announcement list of new publications to be issued in the series: National Bureau of Standards Special Publication 500-.

Name _____

Company _____

Address _____

City _____ State _____ Zip Code _____

(Notification key N-503)

NBS *Technical Publications*

Periodical

Journal of Research—The Journal of Research of the National Bureau of Standards reports NBS research and development in those disciplines of the physical and engineering sciences in which the Bureau is active. These include physics, chemistry, engineering, mathematics, and computer sciences. Papers cover a broad range of subjects, with major emphasis on measurement methodology and the basic technology underlying standardization. Also included from time to time are survey articles on topics closely related to the Bureau's technical and scientific programs. As a special service to subscribers each issue contains complete citations to all recent Bureau publications in both NBS and non-NBS media. Issued six times a year.

Nonperiodicals

Monographs—Major contributions to the technical literature on various subjects related to the Bureau's scientific and technical activities.

Handbooks—Recommended codes of engineering and industrial practice (including safety codes) developed in cooperation with interested industries, professional organizations, and regulatory bodies.

Special Publications—Include proceedings of conferences sponsored by NBS, NBS annual reports, and other special publications appropriate to this grouping such as wall charts, pocket cards, and bibliographies.

Applied Mathematics Series—Mathematical tables, manuals, and studies of special interest to physicists, engineers, chemists, biologists, mathematicians, computer programmers, and others engaged in scientific and technical work.

National Standard Reference Data Series—Provides quantitative data on the physical and chemical properties of materials, compiled from the world's literature and critically evaluated. Developed under a worldwide program coordinated by NBS under the authority of the National Standard Data Act (Public Law 90-396).

NOTE: The Journal of Physical and Chemical Reference Data (JPCRD) is published quarterly for NBS by the American Chemical Society (ACS) and the American Institute of Physics (AIP). Subscriptions, reprints, and supplements are available from ACS, 1155 Sixteenth St., NW, Washington, DC 20056.

Building Science Series—Disseminates technical information developed at the Bureau on building materials, components, systems, and whole structures. The series presents research results, test methods, and performance criteria related to the structural and environmental functions and the durability and safety characteristics of building elements and systems.

Technical Notes—Studies or reports which are complete in themselves but restrictive in their treatment of a subject. Analogous to monographs but not so comprehensive in scope or definitive in treatment of the subject area. Often serve as a vehicle for final reports of work performed at NBS under the sponsorship of other government agencies.

Voluntary Product Standards—Developed under procedures published by the Department of Commerce in Part 10, Title 15, of the Code of Federal Regulations. The standards establish nationally recognized requirements for products, and provide all concerned interests with a basis for common understanding of the characteristics of the products. NBS administers this program as a supplement to the activities of the private sector standardizing organizations.

Consumer Information Series—Practical information, based on NBS research and experience, covering areas of interest to the consumer. Easily understandable language and illustrations provide useful background knowledge for shopping in today's technological marketplace.

Order the above NBS publications from: Superintendent of Documents, Government Printing Office, Washington, DC 20402.

Order the following NBS publications—FIPS and NBSIR's—from the National Technical Information Service, Springfield, VA 22161.

Federal Information Processing Standards Publications (FIPS PUB)—Publications in this series collectively constitute the Federal Information Processing Standards Register. The Register serves as the official source of information in the Federal Government regarding standards issued by NBS pursuant to the Federal Property and Administrative Services Act of 1949 as amended, Public Law 89-306 (79 Stat. 1127), and as implemented by Executive Order 11717 (38 FR 12315, dated May 11, 1973) and Part 6 of Title 15 CFR (Code of Federal Regulations).

NBS Interagency Reports (NBSIR)—A special series of interim or final reports on work performed by NBS for outside sponsors (both government and non-government). In general, initial distribution is handled by the sponsor; public distribution is by the National Technical Information Service, Springfield, VA 22161, in paper copy or microfiche form.

U.S. Department of Commerce
National Bureau of Standards
Gaithersburg, MD 20899

Official Business
Penalty for Private Use \$300